ADVANCES IN INTELLIGENT AND *SOFT COMPUTING* 82

Bing-yuan Cao
Guo-jun Wang
Shui-li Chen
Si-zong Guo (Eds.)

Quantitative Logic and Soft Computing 2010

Springer

Advances in Intelligent and
Soft Computing 82

Editor-in-Chief: J. Kacprzyk

Advances in Intelligent and Soft Computing

Editor-in-Chief

Prof. Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6
01-447 Warsaw
Poland
E-mail: kacprzyk@ibspan.waw.pl

Further volumes of this series can be found on our homepage: springer.com

Vol. 68. V.-N. Huynh, Y. Nakamori, J. Lawry, M. Inuiguchi (Eds.)
Integrated Uncertainty Management and Applications, 2010
ISBN 978-3-642-11959-0

Vol. 69. E. Piętka and J. Kawa (Eds.)
Information Technologies in Biomedicine, 2010
ISBN 978-3-642-13104-2

Vol. 70. Y. Demazeau, F. Dignum, J.M. Corchado, J. Bajo Pérez (Eds.)
Advances in Practical Applications of Agents and Multiagent Systems, 2010
ISBN 978-3-642-12383-2

Vol. 71. Y. Demazeau, F. Dignum, J.M. Corchado, J. Bajo, R. Corchuelo, E. Corchado, F. Fernández-Riverola, V.J. Julián, P. Pawlewski, A. Campbell (Eds.)
Trends in Practical Applications of Agents and Multiagent Systems, 2010
ISBN 978-3-642-12432-7

Vol. 72. J.C. Augusto, J.M. Corchado, P. Novais, C. Analide (Eds.)
Ambient Intelligence and Future Trends, 2010
ISBN 978-3-642-13267-4

Vol. 73. J.M. Corchado, P. Novais, C. Analide, J. Sedano (Eds.)
Soft Computing Models in Industrial and Environmental Applications, 5th International Workshop (SOCO 2010), 2010
ISBN 978-3-642-13160-8

Vol. 74. M.P. Rocha, F.F. Riverola, H. Shatkay, J.M. Corchado (Eds.)
Advances in Bioinformatics, 2010
ISBN 978-3-642-13213-1

Vol. 75. X.Z. Gao, A. Gaspar-Cunha, M. Köppen, G. Schaefer, and J. Wang (Eds.)
Soft Computing in Industrial Applications, 2010
ISBN 978-3-642-11281-2

Vol. 76. T. Bastiaens, U. Baumöl, and B.J. Krämer (Eds.)
On Collective Intelligence, 2010
ISBN 978-3-642-14480-6

Vol. 77. C. Borgelt, G. González-Rodríguez, W. Trutschnig, M.A. Lubiano, M.Á. Gil, P. Grzegorzewski, and O. Hryniewicz (Eds.)
Combining Soft Computing and Statistical Methods in Data Analysis, 2010
ISBN 978-3-642-14745-6

Vol. 78. B.-Y. Cao, G.-J. Wang, S.-Z. Guo, and S.-L. Chen (Eds.)
Fuzzy Information and Engineering 2010
ISBN 978-3-642-14879-8

Vol. 79. A.P. de Leon F. de Carvalho, S. Rodríguez-González, J.F. De Paz Santana, and J.M. Corchado Rodríguez (Eds.)
Distributed Computing and Artificial Intelligence, 2010
ISBN 978-3-642-14882-8

Vol. 80. N.T. Nguyen, A. Zgrzywa, and A. Czyzewski (Eds.)
Advances in Multimedia and Network Information System Technologies, 2010
ISBN 978-3-642-14988-7

Vol. 81. J. Düh, H. Hufnagl, E. Juritsch, R. Pfliegl, H.-K. Schimany, and Hans Schönegger (Eds.)
Data and Mobility, 2010
ISBN 978-3-642-15502-4

Vol. 82. B.-Y. Cao, G.-J. Wang, S.-L. Chen, and S.-Z. Guo (Eds.)
Quantitative Logic and Soft Computing 2010
ISBN 978-3-642-15659-5

Bing-yuan Cao, Guo-jun Wang, Shui-li Chen,
and Si-zong Guo (Eds.)

Quantitative Logic and Soft Computing 2010

Volume 2

Springer

Editors

Prof. Bing-yuan Cao
School of Mathematics and Information Science
Guangzhou University
230 Wai Huan Xi Road,
Guangzhou Higher Education Mega Center, ZIP 510006,
Guangdong, China
E-mail: Caobingy@163.com

Prof. Guo-jun Wang
Department of Mathematics,
Shaanxi Normal University,
Xi'an 710062, China
E-mail: gjwang@snnu.edu.cn

Prof. Shui-li Chen
School of Science,
Jimei University,
Xiamen, Fujian 361021,
China
E-mail: Sgzx@jmu.edu.cn

Prof. Si-zong Guo
Institute of Mathematics and Systems Science,
Liaoning Technical University
123000, China
E-mail: lnzxtaw32@126.com

ISBN 978-3-642-15659-5 e-ISBN 978-3-642-15660-1

DOI 10.1007/978-3-642-15660-1

Advances in Intelligent and Soft Computing ISSN 1867-5662

Library of Congress Control Number: 2010933975

Typeset & Cover Design: Scientific Publishing Services Pvt. Ltd., Chennai, India.

Printed on acid-free paper
5 4 3 2 1 0
springer.com

Foreword

Admittedly, the notion “intelligence or intelligent computing” has been around us for several decades, implicitly indicating any non-conventional methods of solving complex system problems such as expert systems and intelligent control techniques that mimic human skill and replace human operators for automation. Various kinds of intelligent methods have been suggested, phenomenological or ontological, and we have been witnessing quite successful applications. On the other hand, “Soft Computing Techniques” is the concept coined by Lotfi Zadeh, referring to “a set of approaches of computing which parallels the remarkable ability of the human mind to reason and learn in an environment of uncertainty, imprecision and partial truth.” Such a notion is well contrasted with the conventional binary logic based hard computing and has been effectively utilized with the guiding principle of “exploiting the tolerance for uncertainty, imprecision and partial truth to achieve tractability, robustness and low solution cost.” The soft computing techniques are often employed as the technical entities in a tool box with tools being FL, ANN, Rough Set, GA etc. Based on one’s intuition and experience, an engineer can build and realize human-like systems by smartly mixing proper technical tools effectively and efficiently in a wide range of fields. For some time, the soft computing techniques are also referred to as intelligent computing tools.

Though these intelligent and soft computing techniques have been found to be very effective in describing and handling relatively simple human-related systems, we find that the existing theories and related conceptual foundations on soft and intelligent techniques need to be further improved and advanced to cope with the very complex nature of “human,” the crux of the human in the loop system. We observe that the characteristics of “human”, in view of human robot interaction, for example, is time varying, inconsistent, high dimensional, susceptive to noise, ambiguous, subjective and local. In particular, we find that the notion of “approximation” needs to be made more precise as we use those soft computing tools in a mixed way. One of the powerful directions is suggested by those scholars working on “quantitative logic.” As a logic that is more mathematical than verbal, the quantitative logic combines the mathematical logic and the probability computation in a way to provide a graded approach to many-valued propositional logic and predicate logic systems as well explained in a seminal paper by G. J. Wang and H. J. Zhou in Information Sciences. I am very glad to learn that this edited volume is intended to include research results on quantitative logic and other theoretical studies on soft computing, especially including several important theoretical expeditions by renowned scholars such as G. J.Wang and H. J. Zhou, E. P. Klement, R. Mesiar and E. Pap, V. Novak, I. Perfilieva, W. K. Ho, D. S. Zhao and W. S. Wee, Paul P. Wang and C. H. Hsieh, etc.

It has been my honor to review of the volume and comment on this important contribution to the area of intelligent and soft computing. I am very sure that this edition will be a valuable addition to your list of references.

Z. Zenn Bien, Ph.D.
Chaired Professor
School of Electrical and Computer Engineering
UNIST, Korea

Preface

This book is the proceedings of the 2nd International Conference on Quantitative Logic and Soft Computing (QL & SC 2010) from Oct. 22-25, 2010 in Xiamen, China. The conference proceedings is published by Springer-Verlag (Advances in Intelligent and Soft Computing, ISSN: 1867-5662).

This year, we have received more than 165 submissions. Each paper has undergone a rigorous review process. Only high-quality papers are included. The 2nd International Conference on Quantitative Logic and Soft Computing (QL & SC 2010), built on the success of previous conferences, the QL & QS 2009 (Shanghai, China), is a major symposium for scientists, engineers and practitioners in China to present their updated results, ideas, developments and applications in all areas of quantitative logic and soft computing. It aims to strengthen relations between industry research laboratories and universities, and to create a primary symposium for world scientists in quantitative logic and soft computing fields as follows:

1) Quantitative Logic.
2) Fuzzy Sets and Systems.
3) Soft Computing.

This book contains 83 papers, divided into five main parts:

In Section I, we have 7 papers on "Keynote Speakers".

In Section II, we have 24 papers on "Quantitative Logic".

In Section III, we have 25 papers on "Fuzzy Sets and Systems".

In Section IV, we have 27 papers on "Soft Computing".

In addition to the large number of submissions, we are blessed with the presence of seven renowned keynote speakers and several distinguished panelists and we shall organize workshops.

On behalf of the Organizing Committee, we appreciate Jimei University and Shanxi Normal University in China, and International Fuzzy Mathematics Institute in USA. We are grateful to the supports coming from the international magazines published by Springer-Verlag GmbH. We are showing gratitude to the members of the Organizing Committee and the Program Committee for their hard work. We wish to express our heartfelt appreciation to the keynote and panel speakers, workshop organizers, session chairs, reviewers, and students. In particular, we are thankful to the sponsors of the conference, namely, Xiamen Ropeok Science and Technology Co., Ltd, Xiamen Smart Technology Development Co., Ltd, Xiamen Feihua Environment Protective Material Co., Ltd., ChinaSoft International Limited and Pay Seng Koon Mathematics Science Activities Fund. Meanwhile, we thank the publisher, Springer, for publishing the QL & SC 2010 proceedings as J. Advances in Intelligent and Soft Computing (AISC). Finally, we appreciated all the authors and participants for their great contributions that made this conference possible and all the hard work worthwhile.

October 2010
Xiamen, P.R. China

Bing-yuan Cao
Guo-jun Wang
Shui-li Chen
Si-zong Guo

Organization

Organizing Committee

Conference Chairs: Ru-qian Lu (China) Guo-jun Wang (China)
Honorary Chairs: Z. Bien (Korea)

Program Committee

Chair: Paul P. Wang (USA)
Co-chair: Shui-li Chen (China) Yi-xiang Chen (China)

Members

Bing-yuan Cao (China) Pei-chann Chang (Taiwan) Guo-long Chen (China)
Ji-shu Cheng (China) Francesc Esteva (Spain) Jiu-lun Fan (China)
M. Fedrizzi (Italy) Tai-he Fan (China) S. Gottwald (Germany)
Jin-wu Gao (China) Si-cong Guo (China) Cheng-ming Hu (USA)
Shi-er Ju (China) E.E. Kerre (Belgium) K. H. Kim (USA)
Hong-xing Li (China) Jin-jun Li (China) Qing-guo Li (China)
Shao-zi Li (China) Yong-mei Li (China) Yong-ming Li (China)
Bao-ding Liu (China) Min Liu (China) San-yang Liu (China)
Zhi-qiang Liu (China) Masao Mukaidono (Japan) Kyung Hwan Oh (Korea)
Dao-wu Pei (China) Irina Perfilieva (Czech Republic) D.A. Ralescu (USA)
Mehdi Roopaei (Iran) Da Ruan (Belgium) Kai-quan Shi (China)
Yan Shi (Japan) Mei-qing Wang (China) Zhen-yuan Wang (USA)
Wei-zhi Wu (China) Luo-shan Xu (China) Yang Xu (China)
Zhong-qiang Yang (China) Yi-yu Yao (Canada) Bin Zhao (China)
Dong-sheng Zhao (Singapore) Xi-shun Zhao (China)
Cheng-yi Zhang (China) Guo-qiang Zhang (USA)
Guang-quan Zhang (Australia) Chang-le Zhou (China) Wu-neng Zhou (China)

Local Arrangements Chair

Yun-dong Wu (China)
Co-chairs : Guo-rong Cai (China) Zhen-kun Huang (China)
Secretaries : Jia-liang Xie (China) Jun-fa Li(China)
Members : Xiang-qun Shen (China) Jun Yan (China)

Publicity Chair : Su-jun Weng (China)
Co-chairs : Xiao-hui Liu (China) Dong-xiao Zhang (China)
Member : Jia-xiang Chen (China)

Publication Chairs : Bing-yuan Cao (China) Guo-jun Wang (China)
Co-chairs : Shui-li Chen (China) Si-zong Guo (China)
Member : Pei-zi Chen (China)

Contents

Keynote Speakers

Quantitative Logic

Fuzzy Sets and Systems

Soft Computing

Learning Structure of Human Behavior Patterns in a Smart Home System

Zeungnam Bien[1] and Sang Wan Lee[2]

[1] School of Electrical and Computer Engineering,
Ulsan National Institute of Science and Technology,
Ulsan 689-805, Korea
e-mail: zbien@unist.ac.kr
[2] Neuro System Research Group,
Department of Bio and Brain Engineering,
Korea Advanced Institute of Science and Technology,
Daejeon 305-701, Korea
e-mail: bigbean0@gmail.com

Abstract. A growing proportion of the aged in population provokes shortage of caregivers and restructuring of living spaces. One of the most promising solutions is to provide with a smart home environment that ensures independence of users. In this paper, we first call attention to the fact that a learning capability of human behavior patterns can play a central role in adequate functioning of such systems. Specifically, we give an overview of important related studies to illustrate how a variety of learning functions can be successfully incorporated into the smart home environment. We then present our approaches towards the issues of life-long learning and non-supervised learning, which are considered essential aspects of a smart home system. The two learning schemes are shown to be satisfactory in facilitating independent living over different time scales and with less human intervention. Finally, we mention about a prospective model of a future smart home.

Keywords: learning, life-long learning, non-supervised learning, human behavior patterns, smart home.

1 Introduction

According to the statistical bureau reports of Asian, European, and US [1-4], the rate of the aged population has dramatically increased. For example, the expected rate amounts to more than 25 percent in Korea, Japan, and Germany, and 20 percent in US, England, and France. The proportion of the aged is growing worldwide, and it is expected that this will be tripled by 2050 [5]. We are concerned about this situation because this may unfold shortage of caregivers and living spaces. Note that the elderly suffers from the problem of degenerated motor functions which lead social isolation with affective disorders. To be more specific, the

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 1–15.
springerlink.com

proportion of the aged in Korea is expected to be tripled by 2030, whereas the social capability to support them will be doubled [2]. The world population bulletin has reported that "Old-age dependency ratio" (number of people 65 or older over number of people ages 20 to 64) is expected to be doubled by 2045 [5]. As a result, the percentage of the worldwide population over the age 65 will be doubled within three decades; 7.4% in the year 2005 will become 15.2% in the year 2045.

It appears from the population reports that our society will suffer from the lack of young people who are capable of supporting older people. This problem has long been tackled by many social groups as well as researchers in the field of robotics. As a feasible solution of particular interest, the robotic approaches have been coupled with the smart home technologies so as to successfully substitute human caregivers with some automated service agents.

Table 1. List of Well-known Smart Homes

	Name	Nation	Level of Intelligence	Web/ related works
1	Adaptive House	USA	3	(1)/[7]
2	Aware Home	USA	1	(2)
3	Ceit Living Lab	Austria	1	(3)
4	Cogknow	Europe-wide	1	(4)
5	community computing	Korea	1	(5)/[8]
6	Context visualization	Korea	1	-/[9]
7	Context-aware unified remocon	Korea	1	(6)/[10]
8	DLF SmartHome	UK	1	(7)
9	Domotics	EU	2	(8)/[11]
10	DOMUS	Canada	1	(9)/[12]
11	Easyliving Room	USA	1	-/[13]
12	Futurelife Haus	Swiss	1	(10)
13	Gator Tech Smart House	USA	1	(11)/[14]
14	Global Village initiative	Worldwide	1	(12)
15	Home Control Center	Finland	1	(13)
16	Home Depot Smart Home	USA	3	(14)
17	Home Network	Korea	1	(15)/[15]
18	Human Space	Korea	1	(16)
19	ICT-ADI	Worldwide	3	(17)
20	In-HAM Home Lab	Belgium	1	(18)
21	Inhaus-Zentrum in Duisburg	Germany	1	(19)
22	Intelligent Sweet Home	Korea	3	(20)/[16]
23	Intelligent Workplace	USA	1	(21)
24	IR remocon module-based Smart Home	Korea	1	-/[17]
25	IR-based User Detection System	Korea	1	-/[18]
26	Kompetenzzentrum Smart Environments	Germany	1	(22)
27	Living Tomorrow in Amsterdam	Belgium	1	(23)
28	MARC smart home	USA	1	(24)/[19]
29	Microsoft Home	USA	1	(25)
30	Millennium Homes	UK	1	(26)/[20]
31	MIT house_n	USA	1	(27)
32	MIT smart city	USA	1	(28)
33	NUADU	Europewide	1	(29)

Table 1. (*continued*)

34	OSGi based Intelligent Home	Korea	1	-/[21]
35	Robotic Room	Japan	3	(30)
36	Sentient Computing	UK	1	(31)/[22]
37	SerCHo Showroom	Germany	1	(32)
38	Service Differentiation	Korea	1	(33)/[23]
39	Smart Medical Home	USA	1	(34)
40	Smartest Home	Netherlands	2	(35)
41	SmartHOME	Germany	1	(36)
42	STARhome	Singapore	2	(37)
43	Steward Robot	Korea	1	(38)/[24]
44	Telegerontology	Spain	1	(39)
45	The Intelligent Dormitory 2 (iSpace2)	UK	3	(40)/[25]
46	The Smart SoftWareHouse	Swiss	3	(41)/[26]
47	Toyota Dream House PAPI	Japan	1	(42)
48	TRON Intelligent House	Japan	1	(42)
49	ubiHome	Korea	2	(43)/[27]
50	Welfare Techno House	Japan	2	(44)/[28]
51	ZUMA	USA	1	(45)

(The names are listed in the alphabetical order. The references and the website address are indicated by brackets and parentheses, respectively. The website addresses are listed in the Appendix. Each smart home is labeled with 'level of intelligence' that indicates whether or not equipped with computational intelligence and/or learning capability for human behavior patterns.)

We have examined many smart homes available in the literature/web and have listed in Table 1 those studies on well-developed smart homes. In our survey study, we have noted with special interest that its realization with advanced intelligence has become an issue of common interest worldwide and that the learning capability is considered an important function in many recent developments of home for the elderly, as opposed to an initial concept that mere addition and arrangement of various smart gadgets and devices for home would suffice. In order to pursue this perspective, we take a notion of "level of intelligence" into consideration. The level 1 refers to the case where no particular techniques of computational intelligence are considered in the home development, with the level 2 referring to the case where some computational intelligence techniques are introduced for understanding of human behavior patterns, and the level 3 indicates the case where some form of learning or adaptation for human behavior patterns is incorporated. From the table, it can be seen that the majority of the systems (i.e., 39 out of 51 cases) do not possess any notable function of intelligence; they are merely designed to react to the user's request or to execute a preprogrammed procedure. On the other hand, we find only seven smart homes have been implemented with some functions of adaptation or learning of the user's behavior patterns. Recall that most of the contents of the proactive services require understanding of human minds and intention, some of which are possible through observation and analysis of human behaviors and gestures. Therefore, in this paper we have put a great deal of emphasis on realization of the learning capability for human behavior patterns as the most important factor of the automated service agents in a smart home.

This paper thus begins with an overview of studies that successfully cover the top-ic of human behavior pattern learning (Section 2). We then present in Section 3 our two approaches toward the issues of life-long learning and non-supervised action sequence learning which we think are central for a smart home with a proactive service capability. We then show that these capabilities are satisfactory to facilitate independent living over a variety of time scales and with less human intervention. Finally, we give several remarks on a prospective smart home under consideration at UNIST.

2 Understanding of Human Behavior Patterns in a Smart Home Environment

According to an up-to-date brain theory based on the free energy minimization principle [29], the first task in a brain's work flow is to build up a recognition model, i.e., to obtain a probability distribution of 'causes'[30] given a set of observations.

The same principle can be applied to the smart home system. The first step towards learning is to recognize human behavior patterns, and then know their 'causes'. It is a necessary process because various activities of a user give rise to a multitude of changes in the environment; for example, one of the changes could be a change in the position of an object, a change in temperature, or even a change in a state of human subject. They are called 'evidences' or observations. In a smart home environment, what the home management system would do first is to collect these evidences, which can be thought of as 'samples' from the system's point of view. All these samples that the system has as evidences are considered to arise from putative causes. The next step is then to get a model of causes, which is often referred to as an inference process, a recognizer, or a classifier. An overview on these two processes is given next.

2.1 Activity Recognition and Its Applications

A typical operational task of a smart home is to simply detect or recognize activities of users to provide a proactive service or to monitor them for the purpose of health care. In [31], the location information of a user accumulated over time is used for estimation. More examples deal with the issue of activity monitoring. In [31], a data mining technique has been applied for detection of frequent trajectory activity patterns. In [32], a clustering analysis of human activity patterns is done to detect, visualize, and interpret human activities. In [33], a Bayesian network discovers relationships between sensor signals and observed human activities. Note that recognition of a user's activity patterns can be applied to various systems, including an adaptive phone profile control system [34] or a system that characterizes behavioral patterns of Alzheimer's disease patients [35].

The fore-mentioned activity recognition capability serves to predict the user's next actions [36]. For example, in [37] is presented a prediction system of requested resources by learning the collective resource allocations and navigation patterns of users. Moreover, it can be easily conceived that information of the user's activity patterns can incorporate an individual preference or the circumstances. In the concept "Ubiquitous Robotic Companion"[38,39], the system has been designed to recommend appropriate services by taking the circumstances and a user preference into account.

Another line of studies is to utilize the patterns of a user's behavior to augment some functions of E-learning systems; this study reveals that an interactive and data-driven approach is crucial for effective education. In the proposed education supporter [40], a simple rule is used to analyze habits of learning. In [41], an adaptive learning sequence prediction method has been applied to a ubiquitous e-learning system. In [42], throughout a dynamic assessment the authors have found multiple features that are believed to motivate learners. In addition, adaptation and lifelong learning issues have been raised. For example, in [43], the system provides an adaptive learning contents based on learner's preference and contextual environment. Kay has presented a lifelong user model [44] that can support personalized lifelong learning.

2.2 Understanding of Irregular Behavior Patterns

Performance of an activity recognition system is often unsatisfactory, for which many argue that the origin of this problematic situation lies in irregularity of human behavior patterns; for this reason, many researchers have suspected that applications of conventional rule bases or statistical models may not cope with this difficulty. In support of this claim, we give an overview of those studies that deal with putative causes of irregularity in human behavior patterns.

Two factors of uncertainty and emotion are paid particular attention for the irregularity issue. A study in [45] deals with the problem of learning and predicting inhabitants' activities under uncertainty and the presence of incomplete data. Another study has asserted that an emotion is an utterly important factor that influences user's decision making process [46]; this perspective brings about a successful recommender system.

The second approach focuses on understanding of exceptional activities, as exemplified in [47]. In [48], a model capable of detecting abnormality in user behavior patterns has been proposed and applied to an access card system. In [49], a learning and prediction model of health vitals of a resident has been suggested; the proposed model utilized information of exceptional activities to provide a precaution to the residents. In [50], a comprehensive analysis has been conducted using 22 users' activity recordings in an assisted living environment. By modeling circadian activity rhythms and observing the deviations in those patterns, the proposed system provides a successful warning service for caregivers.

3 Life-Long and Non-supervised Learning of Human Behavior Patterns in a Smart Home Environment

We find that most of the previous studies illustrated in Section 2 do not take two important issues into consideration: (1) learning over a variety of time scales (i.e., time scalability) and (2) learning without intervention (i.e., unsupervised learning capability) [51]. In response to these problems, we have suggested a Life-long learning framework [52] and a non-supervised learning framework [53], which is intended to adaptively incorporate changing patterns of a user over long time and provide a fully automated process of learning sequential actions, respectively.

3.1 Life-Long Learning of Human Behavior Patterns

Since a user's behavior patterns tend to change over time, a system built and implemented a long while ago would not perform satisfactorily in the future time. One possible solution may be preserving previously learned knowledge as long as it does not contradict the current task [54]. A particular concern is, however, that we do not know how fast/slow or how frequently the changes occur. In order for the system to learn and adapt to a user's behavior patterns that change over time, one is required to deal with time-varying and non-stationary targets [52] such as bio signals, gestures, or moving motions.

The proposed learning framework has been carefully designed in a way that the inductive and the deductive learning process are streamlined. This yields the two-way information flow diagram as shown in Fig. 1 in which one process engages with the other via PFRB (Probabilistic Fuzzy Rule Base [55]) and CDS (Context Description Set); PFRB refers to the learning model with parameterized rules, and CDS refers to the context of an environment which might change over time. In the inductive learning process, the parameters of PFRB and CDS are updated inductively by pairs of I/O data and by environment data. In the deductive learning process, the system prunes some of the rules that are determined to be unnecessary and grafts new rules to reflect a change in an environment.

Interestingly enough, one can find some analogy between the proposed process and a unified brain theory (recently published in Nature [29]) that accounts for human actions, perception, and learning. In the free-energy principle, an iterative optimization occurs throughout a two-way process that includes development of recognition model by a bottom-up process [56] and development of a generative

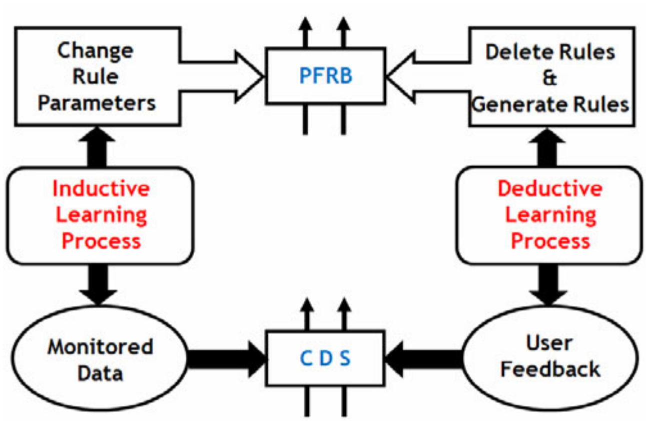


Fig. 1. Mechanism of Learning Process [52]

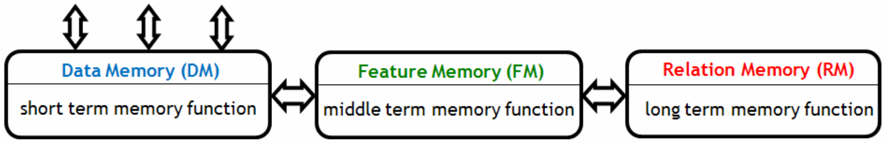


Fig. 2. Memory Subsystem Architecture [52]

model by a top-down process [56]. This can be likened to a two-way process with inductive learning and deductive learning for the following two reasons: (i) Optimization of one model is conducted given the other model iteratively, and (ii) Deductive learning brings about models of 'causes' and their parameters are optimized in the inductive learning process.

Another interesting characteristic of the proposed learning framework is inclusion of memory subsystems. This consists of three units hierarchically: (i) DM (Data Memory) with a STM (Short-Term Memory) function, (ii) FM (Feature Memory) with a mid-term memory function, and (iii) RM (Relation Memory) with a LTM (Long-Term Memory) function. The very first place where the incoming information is stored is DM with STM. Next, neuronal activities (or say data in FM) may arise from the information in STM and finally be passed to the deeper layer called RM. This chain process is likened to the role of Hippocampus in human brain [57].

It is interesting to note by analogy that the proposed learning framework seems well to reflect possible action flow of some memory processes in a human brain.

3.2 Non-supervised Learning of Human Behavior Patterns

Human behavior patterns can be very complex [58-60] in the sense that the corresponding learning process often could be hampered by unplanned or unscheduled actions. Moreover, an automatic discovery of those patterns is required to give a support to people with disabilities or with mental problems such as mild dementia.

The proposed non-supervised learning framework [53] resolves these problems by integrating agglomerative clustering technique and uncertainty-based sequence learning method. The algorithm first starts to discover a meaningful structure of behavior pattern data. In particular, it finds local clusters and then finds optimal partitions of clusters such that they maximize the proposed cluster validity index. Second, by applying Fuzzy-state Q-learning algorithm (FSQL) the system learns action sequences on the basis of the partition sets discovered in the previous process.

Fig. 3 illustrates how the system learns and process some action patterns of a user. We provide a quote from our paper published in [53]:

***Example* [53].** *Referring to Fig. 3, suppose that our task is to program a service robot that takes care of a user named "George". If the states that the robot observes can be defined by "drinking coffee", "watching TV", and "having a meal", then the corresponding action would be interpreted as "serving coffee", "turning on TV", and "serving a meal", respectively. Next, suppose that each state lies in the space of time and temperature; then George's patterns of behavior can be described in the two dimensional space. By applying AIBFC, we get categorized action sets of George's behavioral pattern, such as "When cold morning, George usually drinks coffee.", "When cold afternoon, George likes to watch television.", "When hot morning, George likes to* *watch television.", and "When hot evening, George usually enjoy a meal." Next, from the habits obtained from George, we are able to make a relation among them. Indeed, this will enable the robot to predict what George wants next, based on the previous actions and the corresponding values of certainty. As a consequence, serving coffee at 15:00 under a temperature of 5 degree Centigrade can be quantified as the linear combination of "serving coffee on cold morning with certainty level of 0.2", "serving coffee on cold afternoon with certainty level of 0.7", "serving coffee on hot morning with certainty level of 0.02", and "serving coffee on hot afternoon with certainty level of 0.08".*

This study partly shares the view of the Helmholtz machine [56] in that it constructs the model of 'cause' [29] (i.e., "recognition model" [56]) given the observations. It can be seen from this perspective that our non-supervised learning framework provides an automatic process of developing a sequence recognition model whose underlying structure is optimal in terms of class-separability.

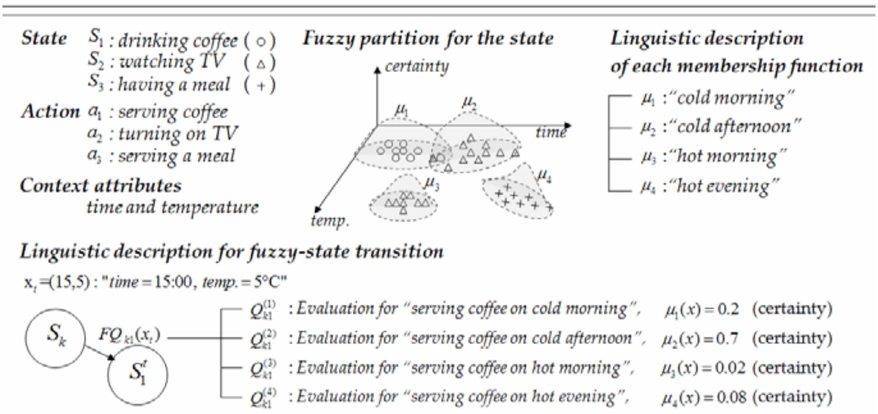


Fig. 3. Example of Action Sequence Prediction Mechanism [53]

In our subsequent study, our learning framework has been applied to a human behavior suggestion system for memory impaired persons [61]. Based on the report that regular daily life can alleviate symptoms of a memory loss, we have introduced a new feature called "averaged frequency" that accounts for how frequently a user tends to take a certain kind of actions. It turns out that our non-supervised learning framework that makes use of the averaged frequency feature successfully improves performance of the human behavior suggestion system.

4 Remarks on a Future Smart Home System for the Aged and the Disabled

It is well-known that a prospective smart home for older persons and persons with physical disabilities should be able to provide human-robot interaction and human-friendly services [6], as mentioned in many studies [62]. The future smart home will include human-centered technologies where important technological components should provide human-friendly interaction with the user [6].

There are several research groups focusing on this issue such as Human Friendly Systems Research Group at Intelligent Systems Institute AIST, Japan and our Human-friendly Welfare-Robotic System Engineering Research Center (HWRS-ERC) at KAIST, Korea. The smart home system at HWRS-ERC (see Fig. 4) provides a Bluetooth-based health monitoring module, a hand-gesture recognition module, a walking assistance, an intelligent wheelchair, a robot coordinator, and a communication server in a way that their cooperation ensure an daily living without caregiver's help.

The main feature we have intended to incorporate into the system is a Human-Robot Interaction (HRI) with human-friendliness (see Fig. 5). Taking the physical difficulties of the aged and the disabled into account, we claim that the two utterly important features of HRI are non-supervised and lifelong learning capability. A non-supervised learning capability can bring less intervention, thereby facilitating autonomous adaptation in 'space'. Equally important is a lifelong learning capability; this enables the system to process information of long-term behavior patterns, thereby facilitating autonomous adaptation in 'time'.

The two approaches of life-long learning and non-supervised learning are expected to be nicely fused in a future smart home. The home system would first undergo a non-supervised learning of human behavior patterns and, as this learning process continues, the accumulated information would be effectively folded by the life-long learning process (a bottom-up process with a blend of inductive learning and deductive learning) while the prediction could be made by unfolding this information (a top-down process). The system that continuously goes through this learning and prediction process would finally develop an internal user model, which we call "robotic clone."

Note that there has been a paradigm shift in a smart home research for the past two decades. In the near future, the notion of 'smart environment' spans a range from a "home" to a "village" or a "city". The project on "global village" [65] serves to illustrate how the system provides large-scale integrated services for dependant people. In this new type of a smart home environment, a living space is

Fig. 4. Intelligent Sweet Home at KAIST [63,64]

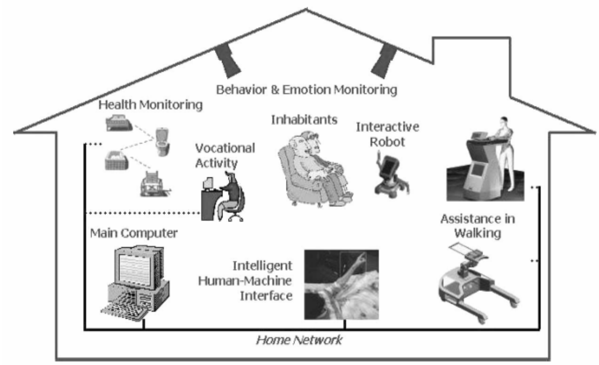


Fig. 5. Smart home with human-friendly service capability [6]

not limited to a single house. On the contrary, it spans a multitude of houses and facilities relevant to the user. This raises rather complex issues, such as seamless information hand-over and adaptation in changing environment. In such complex and dynamic environment, the above-mentioned two types of learning capabilities would have a critical role.

5 Conclusion

In the paper, an overview of smart home studies is given with particular attention to learning functions of human behavior patterns and to the fact that the learning functions can be successfully incorporated into a smart home system. The potential problem in these studies, however, is that most of current smart home systems cannot cope with learning of behavior patterns over a variety of time scales while with minimum human intervention. As effective remedies, a life-long learning framework and a non-supervised learning framework have been suggested. First, the previously suggested life-long learning framework combines the inductive and the deductive learning process, which resemble development of a recognition model via continual top-down and bottom-up processes in a human cortex. Furthermore, the framework becomes more effective by introducing short-term and long-term memory whose presence and functions have been biologically verified. Second, the non-supervised learning framework brings the special type of a recognition model that learns sequential information. Note that the framework is suitable for an automatic service for the elderly and the disabled because the underlying process is gone through with less human intervention.

The future smart environment will span a range from a home to a village. Accordingly, seamless information hand-over and adaptation in changing environment will be pertinent to provision of large-scale integrated services. In such cases, the above-mentioned two types of learning capabilities can provide more effective solutions.

Acknowledgements. This work was supported by the 2009 Research Fund of the UNIST (Ulsan National Institute of Science and Technology).

References

1. Statistical Bureau, Population Reports. The Management and Coordination Agency, Japan (2008)
2. Statistical Bureau, Aged Population Reports, Korea National Statistical Office, Korea (2008)
3. Statistical Bureau, Report of Royal National Institute of Blind People, UK (2008)
4. http://km.paran.com/sub/View.php?k_id=91862&userno=&c_sect=4&c_id=1002004&offset=&DeedChk=N&form=f1&s_cp=reportshop
5. Sanderson, W., Scherbov, S.: Rethinking Age an Aging. Population Bulletin 663 (2008)
6. Stefanov, D.H., Bien, Z., Bang, W.C.: The Smart House for Older Persons and Persons With Physical Disabilities: Structure, Technology Arrangements, and Perspectives. IEEE Transactions on Neural Systems and Rehabilitation Engineering 12 (2004)
7. Mozer, M.C.: The Neural Network House: An Environment that Adapts to its Inhabitants. In: Proceedings of the American Association for Artificial Intelligence Spring Symposium on Intelligent Environments, Menlo Park, pp. 110–114 (1998)
8. You, S.H., et al.: COCOLAB: Supporting Human Life in Ubiquitous Environment by community Computing. In: Proc. Ubiquitous Computing and Network Systems Workshop, JeJu, Korea (2005)
9. Kyoung, K., Seung, H., Hong, J.: A case study on visualization for user's context analysis in home. In: Proc. Korea Human Computer Interaction Conference (2005)
10. Kim, Y., Lee, D.: A personal context-aware universal remote controller for a smart home environment. In: Proc. Advanced Communication Technology, ICACT 2006, The 8th International Conference, vol. 3 (2006)
11. Decamps, E.A., Pecot, F., Royer, G.: Organization of a Domotic Project (Intelligent Building): The Role of Interfaces. Construction Informatics Digital Library, paper w78-1993-34 (1993)
12. Boussemart, B., Giroux, S.: Tangible User Interfaces for Cognitive Assistance. In: 21st International Conference on Advanced Information Networking and Applications Workshops, pp. 852–857 (2007)
13. Brumitt, B., Meyers, B., Krumm, J., Kern, A.: Shafer S. EasyLiving: Technologies for Intelligent Environments. In: Proc. Handheld and Ubiqitous Computing, pp. 12–29 (2000)
14. Helal, S., Bose, R., Chen, C.: The Internal Workings of the Gator Tech Smart House - Middleware and Programming Models. In: Smart Houses: Advanced Technology for Living Independently. Studies in Computational Intelligence. Springer, Heidelberg (2009)
15. Cha, J.: Digital Smart Home. Marine Industry Research Center Report (2008)
16. Lee, S.W., Kim, Y.S., Bien, Z.: Learning Human Behavior Patterns for Proactive Service System: Agglomerative Fuzzy Clustering-based Fuzzy-state Q-learning. In: Proceedings of International Conference on Computational Intelligence for Modeling, Control and Automation (2008)
17. Lee, D.G., Jeong, K.S., Choi, D.J.: Controlling Smart Home based IR-Remote controller. In: Proc. 27th Korea Information Processing Society Spring Conference (2007)

18. Kim, H.H., Ha, K.N., Lee, K.C., Lee, S.: Performance Index for Sensor Arrangement of PIR Sensor-based Indoor Location Aware System. Journal of the Korean Society of Precision Engineering 24, 37–44 (2007)
19. Alwan, M., Dalal, S., Mack, D., Kell, S., Turner, B., Leachtenauer, J., Felder, R.: Impact of Monitoring Technology in Assisted Living: Outcome Pilot. IEEE Transactions on Information Technology in Biomedicine (2007)
20. Sandström, G., Keijer, U.: Integrated Smart Living –Training Flats for Persons with Acquired Brain Dysfunction. Abitare e Anziani Informa 1-2, 85–90 (2003)
21. Park, C.G., Yoo, J.H., Seok, S.H., Park, J.H., Lim, H.I.: Intelligent Home Network Service Management Platform Design Based on OSGi Framework. In: Kim, Y.-T., Takano, M. (eds.) APNOMS 2006. LNCS, vol. 4238, pp. 550–553. Springer, Heidelberg (2006)
22. Addlesee, M., et al.: Implementing a sentient computing system. Computer 34, 50–56 (2001)
23. Choi, J.: Service Differentiation Strategy for Smart Home. Sejong Multimedia Internet Laboratory Internal Report (2003)
24. Park, K.H., Lee, H.E., Kim, Y., Bien, Z.: A Steward Robot for Human-friendly Human-Machine Interaction in a Smart House Environment. IEEE Transactions on Automation Science and Engineering 5, 21–25 (2008)
25. Pounds-Cornish, A., Holmes, A.: The iDorm - A Practical Deployment of Grid Technology. In: 2nd IEEE International Symposium on Cluster Computing and the Grid (2002)
26. Macskassy, S., Hirsh, H., Provost, F., Sankaranarayanan, R., Dhar, V.: Intelligent Information Triage. In: Proc. SIGIR (2007)
27. Hong, D., et al.: Advances in Tangible Interaction and Ubiquitous Virtual Reality. IEEE Pervasive Computing 7, 90–96 (2007)
28. Kawarada, A., Tsukada, A., Sasaki, K.: Automated Mornitoring System For Home Health Care. In: Proceedings of The Fust Joint BMESiEMlS conferens Serving Humanity, Advancing Technology, USA
29. Friston, K.: The free energy principle: a unified brain theory? Nature Reviews Neuroscience 11, 127–138 (2010)
30. Dayan, P., Hinton, G.E., Neal, R.M., Zemel, R.S.: The Helmholtz Machine. Neural Computation 7, 1022–1037 (1995)
31. Manley, E.D., Deogun, J.S.: Location Learning for Smart Homes. In: Proc. 21st International Conference on Advanced Information Networking and Applications Workshops, vol. 2, pp. 787–792 (2007)
32. Zheng, H., Wang, H., Black, N.: Human Activity Detection in Smart Home Environment with Self-Adaptive Neural Networks. In: Proc. IEEE International Conference on Networking, Sensing and Control, pp. 1505–1510 (2008)
33. Albinali, F., Davies, N., Friday, A.: Structural Learning of Activities from Sparse Datasets. In: Proc. Fifth IEEE International Conference on Pervasive Computing and Communications, pp. 221–228 (2007)
34. Valtonen, M., Vainio, A.M., Vanhala, J.: Proactive and adaptive fuzzy profile control for mobile phones. In: Proc. Seventh IEEE International Conference on Pervasive Computing and Communications, pp. 1–3 (2009)
35. Shuai, Z., McClean, S., Scotney, B., Xin, H., Nugent, C., Mulvenna, M.: Decision Support for Alzheimer's Patients in Smart Homes. In: Proc. 21st IEEE International Symposium on Computer-Based Medical Systems, pp. 236–241 (2008)

36. Lu, E.H.C., Tseng, V.S.: Mining Cluster-Based Mobile Sequential Patterns in Location-Based Service Environments. In: Proc. Tenth International Conference on Mobile Data Management: Systems, Services and Middleware, pp. 273–278 (2009)
37. Madkour, A., Sameh, A.: Intelligent Open Spaces: Using Neural Networks for Prediction of Requested Resources in Smart Spaces. In: Proc. 11th IEEE International Conference on Computational Science and Engineering, pp. 132–138 (2008)
38. Moon, A., Kang, T., Kim, H., Kim, H.: A Service Recommendation Using Reinforcement Learning for Network-based Robots in Ubiquitous Computing Environments. In: Proc. The 16th IEEE International Symposium on Robot and Human interactive Communication, pp. 821–826 (2007)
39. Moon, A., Choi, Y., Lee, B.S.: Context-aware user model for personalized services. In: Proc. Third International Conference on Digital Information Management, pp. 858–863 (2008)
40. Dong, M., Ota, K., Cheng, Z., Wang, G.: A Support Method for Improving Learner's Learning Habit Using Behavior Analysis in a Ubiquitous Environment. In: Proc. International Conference on Parallel Processing Workshops, p. 67 (2007)
41. Chang, H., Wang, C., Shih, T.K.: A learning sequencing prediction system for ubiquitous learning based on SCORM sequencing and navigation. In: Proc. First IEEE International Conference on Ubi-Media Computing, pp. 604–609 (2008)
42. Huang, S.H., Wu, T.T., Chu, H.C., Hwang, G.J.: A Decision Tree Approach to Conducting Dynamic Assessment in a Context-Aware Ubiquitous Learning Environment. In: Proc. Fifth IEEE International Conference on Wireless, Mobile, and Ubiquitous Technology in Education, pp. 89–94 (2008)
43. Zhao, X., Ninomiya, T., Anma, F., Okamoto, T.: A context-aware prototype system for adaptive learning content in ubiquitous environment. In: Proc. IEEE International Symposium on IT in Medicine and Education, pp. 164–168 (2008)
44. Kay, J.: Lifelong Learner Modeling for Lifelong Personalized Pervasive Learning. IEEE Transactions on Learning Technologies 1, 215–228 (2008)
45. Shuai, Z., McClean, S., Scotney, B., Nugent, C.: Learning under uncertainty in smart home environments. In: Proc. 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 2083–2086 (2008)
46. Gonzalez, G., De La Rosa, J.L., Montaner, M., Delfin, S.: Embedding Emotional Context in Recommender Systems. In: Proc. IEEE 23rd International Conference on Data Engineering Workshop, pp. 845–852 (2007)
47. Ou, Y., Cao, L., Luo, C., Liu, L.: Mining Exceptional Activity Patterns in Microstructure Data. In: Proc. IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, vol. 1, pp. 884–887 (2008)
48. Leong, A., Fong, S., Yan, Z.: A Logical Model for Detecting Irregular Actions in Physical Access Environment. In: Proc. 18th International Conference on Database and Expert Systems Applications, pp. 560–564 (2007)
49. Xu, J., Maynard-Zhang, P., Chen, J.: Predictive Data Mining to Learn Health Vitals of a Resident in a Smart Home. In: Proc. Seventh IEEE International Conference on Data Mining Workshops, pp. 63–168 (2007)
50. Virone, G., Alwan, M., Dalal, S., Kell, S.W., Turner, B., Stankovic, J.A., Felder, R.: Behavioral Patterns of Older Adults in Assisted Living. IEEE Transactions on Information Technology in Biomedicine 12, 387–398 (2008)
51. Bien, Z.: Fuzzy-based Learning of Human Behavior Patterns. In: Keynote speech in Fuzz-IEEE, Jeju Island, Korea (2009)

52. Bien, Z., Lee, H.R.: Effective learning system techniques for human-robot interaction in service environment. Knowledge-Based Systems 20 (2008)
53. Lee, S.W., Kim, Y.S., Bien, Z.: A Nonsupervised Learning Framework of human Behavior Patterns Based on Sequential Actions. IEEE Transactions on Knowledge and Data Engineering 22 (2010)
54. Grossberg, S.: Nonlinear neural networks: principle, mechanisms and architectures. Neural Networks, 117–161 (2008)
55. Lee, H.E., Park, K.H., Biem, B.: Iterative Fuzzy Clustering Algorithm With Supervision to Construct Probabilistic Fuzzy Rule Base From Numerical Data. IEEE Transactions on Fuzzy Systems 16 (2008)
56. Dayan, P., Abott, L.F.: Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems. MIT Press, Cambridge (2001)
57. Bear, M.F., Connors, B.W., Ma, P.: Neurosicence: Exploring the Brain. Lippincott Williams & Wilkins (2006)
58. Petrovsky, A.V., et al.: Psychology (1986)
59. Burkhardt, R.: Patterns of Behavior. Chicago Univ. Press, Chicago (2005)
60. Lee, W.: Decision Theory and Human Behavior. John Wiley & Sons, Inc., Chichester (1971)
61. Bae, S.H., Lee, S.W., Kim, Y.S., Bien, B.: Fuzzy-State Q-Learning-based Human Behavior Suggestion System in Intelligent Sweet Home. In: Proc. International Conference on Fuzzy Systems (2009)
62. Terano, T.: Fuzzy Engineering Toward Human Friendly Systems. IOS Press, Amsterdam (1992)
63. Bien, Z., et al.: Intelligent Interaction for Human-friendly Service Robot in Smart House Environment. International Journal of Computational Intelligence Systems 1, 78–94 (2008)
64. Park, K.H., et al.: Robotic Smart House to Assist People with Movement Disabilities. Autonomous Robots 22, 183–198 (2007)
65. http://handicom.it-sudparis.eu/gvi/

List of smart home web sites

(1) http://www.cs.colorado.edu/~mozer/nnh/
(2) http://awarehome.imtc.gatech.edu
(3) http://www.ceit.at/333.html
(4) http://www.cogknow.eu/
(5) http://www.cuslab.com/
(6) http://cs.sch.ac.kr
(7) http://www.dlf.org.uk
(8) http://sourceforge.net/projects/stantor/, http://www.domoticspoint.com/
(9) http://domus.usherbrooke.ca/?locale=en
(10) http://www.futurelife.ch
(11) http://www.icta.ufl.edu/gt.htm
(12) http://handicom.it-sudparis.eu/gvi/
(13) http://smarthomepartnering.com/cms/
(14) http://www.smarthome.duke.edu/
(15) http://marine.gnu.ac.kr/
(16) http://www.imedr.org
(17) http://www.ict-asia-france.org/
(18) http://www.in-ham.be
(19) http://www.inhaus-zentrum.de
(20) http://hwrs.kaist.ac.kr/
(21) http://www.cmu.edu/homepage/innovation/2007/spring/intelligent-workplace.shtml
(22) http://www.fokus.fraunhofer.de
(23) http://www.livingtomorrow.com
(24) http://marc.med.virginia.edu/
(25) http://www.microsoft.com/presspass/press/2004/sep04/09-28MSHome10YearPR.mspx
(26) http://people.brunel.ac.uk/~cssrmjp/Site/Projects.html
(27) http://architecture.mit.edu/house_n/
(28) http://cities.media.mit.edu/
(29) http://www.newsfood.com/q/10059/bringing_healthcare_to_peoples_homes_the_nuadu_project/
(30) http://www.ics.t.u-tokyo.ac.jp/index.html
(31) http://www.cl.cam.ac.uk/research/dtg/attarchive/spirit/
(32) http://www.sercho.de
(33) http://milab.sejong.edu/
(34) http://www.futurehealth.rochester.edu/smart_home/
(35) http://www.smart-homes.nl/engels-woning/engels-slimste-woning-homepage/
(36) http://smarthome.unibw-muenchen.de/de/
(37) http://starhome.i2r.a-star.edu.sg/
(38) http://hwrs.kaist.ac.kr/
(39) http://gerontologia.udc.es/
(40) http://iieg.essex.ac.uk/idorm2/index.htm
(41) http://www.ifi.uzh.ch/ddis/research/relational-machine-learning-etc/monitoring-business-activity/
(42) http://tronweb.super-nova.co.jp
(43) http://uvr.gist.ac.kr/
(44) http://www.websciences.org/cftemplate/NAPS/archives/indiv.cfm?ID=20004350
(45) http://www.cs.berkeley.edu/~yurym/

Modeling the Degree of Truthfulness

Paul P. Wang[1] and Chih Hsun Hsieh[2]

[1] Department of Electrical & Computer Engineering,
Pratt School of Engineering Duke University,
Durham, NC 27708, USA
e-mail: ppw@ee.duke.edu
[2] Department of Information Management,
Aletheia University,32, Chen-Li street,
Tamsui 251, Taipei, Taiwan
e-mail: jury@mail.au.edu.tw

Abstract. This paper reports some novel approach on linguistic logic with our intention to realize CWW, Computing With Words, via a simple example which consists of only five words. As a by product, this simple example of the linguistic logical system may serve as a mathematical model which modeling the degree of truthfulness in daily usage. The five words set of a linguistic variable modeling the degree of truthfulness are; true, nearly true, undecided, nearly false and false. We subjectively choose trapezoidal fuzzy numbers as our linguistic truth values in order to model our linguistic logic system. Firstly, some natural operations and linguistic logic operators are defined to suit our objective of developing a closed linguistic variable set. Then the computation of linguistic truth values for this linguistic logical system is developed in order to facilitate us to perform the linguistic inferences. Properties of these natural operations can be derived accordingly. It is perhaps quite rewarding to see numerous linguistic truth relations defined on a single linguistic truth set and linguistic implications ended up with numerous linguistic truth tables. In addition, the linguistic inferences of generalized modus ponens and generalized tollens determined by linguistic compositional rules based on the linguistic truth relation and some natural operations are introduced. The simple examples of the linguistic inferences of the various generalized tautologies are illustrated. Finally, we have proved via a simple dictionary that a closed and self consistent linguistic logical system indeed can be constructed and it is possible to move a chunk of information as modeled by a fuzzy set to a higher level according to the theory of semiotics. These results have shown some promise in realizing the appealing theory of CWW.

Keywords: Linguistic logic; fuzzy number; representation, ordering & ranking of fuzzy sets; natural operators; approximated reasoning; computing with words; fuzzy relation; semiotics; semantics; truthfulness modeling.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 17–38.
springerlink.com

1 Introduction

Truth is a matter of degree, according to Lotfi A. Zadeh [17]. In fuzzy logic, everything is or allowed to be a matter of degree. Furthermore, in fuzzy logic degrees of truth are allowed to be fuzzy. This paper is intended to investigate a somewhat limited notion about the issue of truthfulness and hope to develop some pragmatist theory which has utility value in real life situation. However, the proposed modeling as well as the investigation of the properties of this model may also lead to demonstrate the novel applications of the concept of the 'linguistic variables' and its associated CWW, Computing With Words.

In recent years, we have witnessed a somewhat rapid growth, both in terms of the quality and the quantity of the theories and applications, of the fuzzy methodology and fuzzy logic. On the application front, the consumer electronics products, industrial control processes, bio-medical instrumentations, intelligent information systems, operations research, decision support systems, etc. the progress is quite prominently displayed in front of us. On the theoretical front, we also have observed its impact extended to the basic scientific issues in the sciences of mathematics, physics & chemistry, as well as the biological sciences.

We must say that one important milestone in the technological breakthrough is the development of the concept of the "linguistic variables" modeling by Lotfi A. Zadeh in his 1973 paper [13]. In this paper, Zadeh's intention is to make the use of 'word', modeled by a fuzzy set, as the basic building block from which a sentence, a paragraph and, eventually, a language can be built upon. Fuzzy logic, more general than the set theoretic multi-valued logic, will have linguistic variables assuming certain 'linguistic truth value' and our paper is intended to propose those linguistic truth values in order to begin the task of constructing a specific language capable of modeling the degree of truthfulness concept.

We propose a very simple language L_1 of which there are only five elements, two crispy elements false and true, and three remaining elements represented by fuzzy sets, nearly true, undecided, and nearly false. Even though L_1 is very simple, but this language can easily be generalized, if necessary, for some other applications. We would like, however, to stress that a more generalized Language Lg can be as designated as follows;

$$Lg=\{l_i,\ i=1, 2,..., m\} \text{ where m is a finite positive integer,}$$

and, $l_i \neq l_j$, for all $i \neq j$, l_i belongs to Lg and l_j also belongs to Lg.

The reason for an even very limited language to be very significant can be seen as far back in history as 300 B.C., the master classic "The Hsiao Ching", the most elemental social and religious concept of Chinese people, has been intentionally written using as a small set of words as possible [19]. In fact,it is intended to be used for a five years old child only.

Language Lg is called linguistic logic, LL, because it is an uncertain logical system of which the linguistic truthfulness value itself is a fuzzy set [14]. In [14], the frame work, the shell, of so called CWW, Computing With Words, has been established by Zadeh. In fact, it is also true that CWW is, ultimately, the goal and the realization of the fuzzy logic theory itself. Further more, CWW is necessarily to require the consideration of the truth and the deep understanding of the human cognitive science. As it turns out, this simple example of language L_1 also answers the issue of the modeling of truthfulness from philosophical viewpoint [17].

The acronym LL has been used to represent the nomenclature of linguistic logic, which is an uncertain logical system, where the truth values are fuzzy subsets with unit interval designated by the linguistic labels such as true, nearly true, undecided, nearly false, etc. The study has shown that the linguistic truth set of LL, Lg, can be generated by a context-free grammar [16], with a semantic rule providing a means of computing the meaning of each linguistic truth value in Lg as a fuzzy subset over [0,1] closed interval.

We also observe that LL in general is not closed under the classical logical operations of negation, conjunction, disjunction as well as implication. The result of a natural logical operation on linguistic truth values in Lg would require, in general, a so called linguistic approximation of some linguistic truth value. In fact, Rescher [12] has listed three distinguished features for linguistic logic as follows; (i) The rule of inference whose validity is only approximate rather than being exact. (ii) Linguistic truth values expressed in linguistic terms would necessarily depend upon the semantic meaning associated with the primary truth value such as true or false, as well as their modifiers nearly, about, more or less, etc. (iii) The truth tables now become imprecise truth tables! This is due to the difference in linguistic logic as compared with those of classical logical systems such as Aristotelian logic [10], inductive logic [7], and multiple valued logic with set valued truth-values.

With recent tremendous advance in brain research, it has become very clear that much of the human reasoning is approximate and vague in nature, rather than in precise manner. Approximate reasoning [16] can be viewed as a process of finding the approximated solution of a system of relational assignment equation, mathematically speaking. This process can be equationed as a compositional rule of inference of which modus ponens is only a special case. A characteristic feature of approximate reasoning is the uncertainty and nonuniqueness of consequents of imprecise premises. Simple example of linguistic approximate reasoning is: x is *true*; if x is *true*, then y is *true*; therefore y is *true*, where x, y, and *true* are linguistic words or statements of linguistic logic system. There were considerable and visible research progress on the issue of how to handle so called the compositional rule of inference for the linguistic conditional inference during the decades of 1970's and 1980's [12]. Lotfi Zadeh & George Klir led the discussion, with Mizumoto (1979), Baldwin and Pilsworth (1980), and Hans Zimmermann (1982) all contributed to the discussion by proposing different variations of solving the problem [13].

The breakthrough, or a great leap forward, occurred in 1979 when Zadeh introduced the concept of so-called 'approximate reasoning'. The impact of this introduction has been proven to be very significant in several fronts of handling the vagueness information, among other applications, the decision making as well as industrial process control, to name only a couple here [13]. Importance to solving our problem on hand is the representation of the propositions statements by assigning the fuzzy sets as truth value to linguistic variables. We are fully aware of the rules of inference as in classical logic are based upon various tautologies such as modus ponens, modus tollens, and the hypothetical syllogism. These rules of inference are now generalized within the framework of the linguistic logic in order to introduce the concept of so-called ‘approximate reasoning’ [13].

The main tools of reasoning used in the traditional logic are the tautologies [18], to name one example here, such as the modus ponens. If x and y are propositions (crisp defined), then $(x \wedge (x \Rightarrow y)) \Rightarrow y$ or

Premise	x is *true*
Implication	If x, then y
Conclusion	y is *true*

Lotfi Zadeh further pointed out in 1973 [16] that the classical modus ponens can indeed logically extended to now popular ‘generalized modus ponens’ in the similar way as that of the crisp propositions being generalized to the linguistic logic propositions.

For example, let A, A’, B, B’ be linguistic logic propositions. Then the generalized modus ponens can be expressed by

Premise	A’
Implication	If A, then B
Conclusion	B’

One may translate the above symbolic example to a real world example such as follows [18]:

Premise	*This tomato is very red.*
Implication	If *a tomato is red*, then *the tomato is ripe.*
Conclusion	*This tomato is very ripe.*

In addition, the operation of the classical approximated reasoning may also be extended to the linguistic approximated reasoning in similar manner as that of the modus ponens being extended to the generalized modus ponens. Actually, Lotfi Zadeh also suggested the use of the compositional rule of inference for the linguistic rule of inference as a operational device in 1973 [16]. For the above example of generalized modus ponens, the compositional rule of inference can be obtained by

$$B'(y) = \sup_{x \in A} \min\,[A'(x), \mathrm{R}(x, y)],$$

where for all $y \in$ Y, and R (x, y), $x \in$ X, $y \in$ Y be fuzzy relation in X $\times$ Y.

In this juncture, it is perhaps important to reflect the philosophy of the fuzzy methodology in particular and, for that matter, all the methodologies involving mathematics of uncertainty in general. When we deal with a situation involving vagueness information, we are handling a chunk of data, say, a fuzzy set, and we are operating this chunk of data. Ultimately, the operation of the defuzzification can not be avoided, just a matter of when are you required to do it. It is also worthy to note that there is no unique way to perform such a defuzzification, nor the types of defuzzification options are quite open and sometimes quite subjective. In approximated reasoning operation, the defuzzification process, once accomplished, the results can be very rewarding. In CWW, Computing With Words, the defuzzification operation is also a necessary process in order to achieve, as we shall see, our main goal.

Hsieh proposed [6], in 2008, a minimum function, min, and a maximum function, max, via an ordering function, in order to obtain an ordering number for each linguistic truth value in a linguistic truth set. Hsieh also used the ranking fuzzy numbers method aiming at obtaining minimum linguistic truth value and maximum linguistic truth value between two linguistic truth values. We feel it is a good idea to extend the previous results to designate a new minimum function, MIN, and the new maximum function, MAX, now to the n linguistic truthfulness instead of just two. Furthermore, a new useful ORDER function, for the purpose of getting an order number of the linguistic truth value via a decreasing linguistic truth set, is also proposed in this paper. In addition, a linguistic complement function, CMP, for the objective of finding the complementary linguistic truth value of a linguistic truth value is also presented in this paper. We accomplish this by using the concept of the decreasing linguistic truth set, an increasing linguistic truth set, as well as the ORDER function just mentioned above.

Without losing of any generality, we have decided to choose the trapezoidal fuzzy number for representing linguistic truth values in the course of linguistic logic discussion. We also take the liberty of choosing some specific linguistic truth values in a linguistic truth set. GMIR, Graded Mean Integration Representation method of Chen et al. [2-5] has been adopted as our method of defuzzification in this paper. So far as the GMIR approach is concerned, they use the important degree of each corner point of the support set of a fuzzy number. It is interesting to note that some so called natural operations of the linguistic logic are defined making the use of the above mentioned MIN, MAX, and CMP functions, in order to compute linguistic truth values. We thought this is quite a novelty. Furthermore, a linguistic truth relation determined by a linguistic implication defined on a single linguistic truth set and the IF−THEN natural operation on conditional linguistic proposition such as “If *s*, then *t*” is proposed. Moreover, linguistic approximate reasoning under generalized modus ponens and generalized modus tollens are introduced by using a linguistic truth relation and some natural operations of linguistic logic.

In the following, representing and ranking of linguistic truth values in linguistic logic by using Graded Mean Integration Representation method are discussed. MIN, MAX, ORDER, and CMP functions are introduced. AND, OR, NOT, and IF−THEN natural operations of linguistic logic defined on the above functions of MIN, MAX, ORDER, and CMP are introduced. Some properties of natural operations are derived, and are used for computing with linguistic truth values in second section. In third section, linguistic truth relation (LTR) determined by a linguistic implication is introduced. In fourth section, the compositional rules of linguistic approximate reasoning and its simple examples are considered. Concluding remarks are discussed in final section.

2 Linguistic Logic

This paper is aimed at two goals; to present a model handling the issue of truthfulness which has a great deal of applications, and to advance the research on CWW, of which the principle of semiotics can be instrumental in implementation of the CWW. The centrality of the above two goals, however, is the development of so called 'linguistic logic'. This section in intended to document the details which ultimately leads to discover and to establish the ‘linguistic logic’ as a respective logical system at higher level as compared with the conventional logic system. Surprisingly, at least for a somewhat well constructed and simple minded CWW operations, we are able to build all essential properties about this linguistic logic. Since the discipline of logic is to study the methodology and the principle of reasoning in all possible form [9], conventional logic must deal with the propositions that are either true or false and nothing in between. Consequently, the degree of truthfulness evidently can not be modeled, this is one of the main motivation for this very investigation and, as it turns out, would take the principle of semiotics by moving a chunk of information to a higher level, namely the linguistic logic to accomplish our goal. This is what we shall develop in this section II. In addition, as we shall see, the generalized negation will be much more interesting as we move to a higher plane! In other words, the conventional logic's negation obviously is much more simpler.

The very notion of CWW, inevitably must invoke linguistic principle and natural language theory. No surprise, as we just mentioned, the even more fundamental semiotics theory turns out to be very helpful in our development of this section which ,to say the least, will test the feasibility of much more difficult task required for CWW from the input-output system's viewpoint. We hope this paper is a good first step!

2.1 Representations for Linguistic Truth Value

In the course of developing our proposed theory on modeling with the truthfulness concept, we must invoke another concept of fuzzy number in the domain of the mathematics of uncertainty. Kaufman [8] gives the definition of a fuzzy number that a fuzzy number in R is a fuzzy subset of R that is convex and normal. Thus a fuzzy number can be considered as a generalization of the interval of confidence. Furthermore, a fuzzy number is more or less a subjective datum and it is a valuation, rather then a measure as per probability theory. More developments on the theory of the fuzzy number can be found, as represented by the works of Dubios and Prade in their articles and books [12, 13].

One successful application of the concept of linguistic variable can be found in the work of Carlsson et al. [1], of which they demonstrated the appealing of using linguistic variables in daily management tasks. The real reason behind this success, however, is the ability of this approach can indeed handle the vagueness and uncertainty which exist in so many places. Furthermore, it is the fuzzy set theory which provides a mathematical modeling in order to achieve what we set out to accomplish [19]. A linguistic variable in this paper can be regarded as a variable whose value is a fuzzy number or a variable whose value is defined in linguistic term.

The truth table in Boolean logic or classical logic can be constructed with ease because there are finite numbers of entries. Unfortunately, the truth table can not be constructed in general, unless other constraints are imposed upon [18], for the case of linguistic logic. Theoretically speaking, there are infinitive entries for linguistic logic's truth table of any logical operator. Should we impose the constraint on the truth value to be finite number of the fuzzy subsets such as true, nearly truth, undecided, nearly false and false, then we will be able to show it is possible to construct a truth table for linguistic logic as well. Zadeh has suggested a somewhat simplified model for the issue of the degree of truthfulness in 1973 [14], of which he has chosen a set of linguistic variables true, false, undecided and unknown for the purpose of modeling the degree of the truthfulness. Zadeh's proposal, even though quite appealing from the philosophical point of view, but quite difficult when we try to model using the mathematics of uncertainty. This is the main reason a different set of linguistic variables are proposed here to further the computing with words applications.

Next step of the realization task would be the selection of fuzzy sets for modeling our linguistic variables. We have chosen judiciously the trapezoidal fuzzy numbers through out this paper because they all satisfy the basic characteristics to be a normal fuzzy number.

Let L be the set with m elements of the truthfulness linguistic variables l_i, where m is a positive integer,

$$L = \{l_i, i = 1, 2, \ldots, m\}. \tag{1}$$

In addition l_i is represented by a trapezoidal fuzzy number, and $l_i \neq l_j$, for all $i \neq j$, l_i and $l_j \in L$.

Furthermore, as mentioned above, we have judiciously chosen L_1={*true*, *nearly true*, *undecided*, *nearly false*, *false*}. The five trapezoidal fuzzy sets have their four corners (a_1, a_2, a_3, a_4) as follows: *true* = (1, 1, 1, 1), *nearly true* = (0.6, 0.7, 0.9, 1), *undecided* = (0.3, 0.4, 0.6, 0.7), *nearly false* = (0, 0.1, 0.3, 0.4), *false* = (0, 0, 0, 0). All these five fuzzy sets are shown in Fig.1.

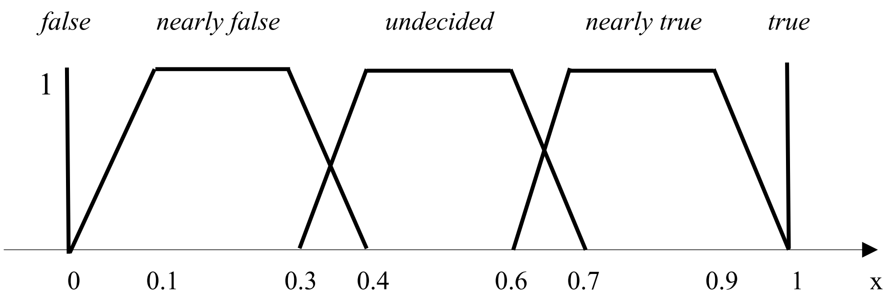


Fig. 1. The linguistic truth values in L_1 are represented as trapezoidal fuzzy numbers.

GMIR, Graded Mean Integration Representation, method was introduced by Chen et al. GMIR is based on the integral value of graded mean h-level of a generalized fuzzy number to represent a fuzzy number. The philosophy behind this proposed method is to adopt the grade as the degree of each point of the support set of a fuzzy number. Suppose l_i is the linguistic value represented by a trapezoidal fuzzy number with corner values (a_1, a_2, a_3, a_4), then according to GMIR method, the representation of l_i, $P(l_i)$ can be evaluated by equation (2).

$$P(l_i) = \frac{a_1 + 2a_2 + 2a_3 + a_4}{6}. \tag{2}$$

Note that equation (2) is consistent for all L_1, including the linguistic truth value 'true' and 'false' with corners values (0,0,0,0) and (1,1,1,1).

2.2 MIN and MAX Operators of Linguistic Truth Values

Since the ordering is an important concept in mathematics, we also need this ordering information in real world application. This has motivated Chen et al to introduce GMIR for ranking fuzzy numbers. Suppose order of l_1 and l_2 are such, $l_1 \geq l_2$, then we can find $P(l_1) \geq P(l_2)$ via equation (2). This allows us to find the ranking of fuzzy numbers, i.e., their ordering using equation (2). Hence, the defuzzification process can then be accomplished [3].

For example, we can get the representation of each linguistic truth value in the above L_1 by equation (2) respectively as follows.

P(*true*) = 1, P(*nearly true*) = 0.8, P(*undecided*) = 0.5, P(*nearly false*) = 0.2, and P(*false*) = 0.

Based on the above results, we can find that the ranking order of linguistic truth values in L_1 as fellows;

true > *nearly true* > *undecided* > *nearly false* > *false*.

In this section, suppose MIN is denoted as a minimum function, in which we can get a minimum linguistic truth value as determined by representations of n linguistic truth values by using equation (2), and be expressed as

$$l_{min} = \text{MIN}\ [l_j, j = 1, 2, \ldots., n], \tag{3}$$

where $l_{min} \in \{l_j, j = 1, 2, \ldots., n\}$ is the minimum linguistic truth value when P (l_{min}) ≤ P (l_j), for all l_j, j = 1, 2, …., n.

For example, the minimum linguistic truth value of four linguistic truth values, *false*, *nearly false*, *true*, and *false*, can be obtained, likewise, as follows;

l_{min} = MIN [*false*, *nearly false*, *true*, *false*]

= *false*.

In addition, in order to find maximum linguistic truth value from n linguistic truth values, a maximum function, MAX, is determined by representations of n linguistic truth values, and be expressed as

$$l_{max} = \text{MAX}\ [l_j, j = 1, 2, \ldots., n], \tag{4}$$

where $l_{max} \in \{l_j, j = 1, 2, \ldots., n\}$ is maximum linguistic truth value when P (l_{max}) ≥ P (l_j), for all l_j, j = 1, 2, …., n.

For example, the maximum linguistic truth value of five linguistic truth values, *nearly true*, *nearly false*, *true*, *false*, and *undecided*, can be found, likewise, as follows;

l_{max} = MAX [*nearly true*, *nearly false*, *true*, *false*, *undecided*]

= *true*.

2.3 Natural Operators for Linguistic Logic

At this point, we introduce three important linguistic logic operators. It is perhaps rewarding to see that the extension of the concept of three operators from classical Boolean Logic to those of the linguistic logic. By doing so, we are extending the known classical results to a whole level up! The flavor of semiotic concept hence is realized via judiciously chosen constraints in this modeling of the truthfulness. It is perhaps interesting to see in the future whether the extension can be even raised up yet to another whole level? In this section, assume that a, b, be two linguistic truth values in *L*. Some natural operators of linguistic logic, linguistic logical operators, are introduced as follows.

2.3.1 Natural Operator AND

Suppose symbol $\wedge$ be denoted to AND natural operation in which it is a binary operation, then AND operation of a and b, a AND b or $a \wedge b$, is given by using the above MIN function, equation (3), as follows

$$a \wedge b = \text{MIN}\ [a, b]. \tag{5}$$

For a, b in L_1, linguistic truth table of $a \wedge b$ can be obtained by equation (5), and the results are summarized in the following Table 1.

Table 1. The linguistic truth table of a AND b ($a \wedge b$)

a \ b	*true*	*nearly true*	*undecided*	*nearly false*	*false*
true	*true*	*nearly true*	*undecided*	*nearly false*	*false*
nearly true	*nearly true*	*nearly true*	*undecided*	*nearly false*	*false*
undecided	*undecided*	*undecided*	*undecided*	*nearly false*	*false*
nearly false	*nearly false*	*nearly false*	*nearly false*	*nearly false*	*false*
false	*false*	*false*	*false*	*false*	*false*

2.3.2 Natural Operator OR

OR linguistic logical operation, linguistic logical union, of two linguistic truth values by using the above MAX function is described in this subsection. Suppose symbol $\vee$ be presented to OR natural operation, then OR natural operation of a and b, a OR b or $a \vee b$, is given by using the above MAX function as follows;

$$a \vee b = \text{MAX}\ [a, b]. \tag{6}$$

For a, b in L_1, linguistic truth table of $a \vee b$ can be obtained by equation (6), and the results are summarized in the following Table 2.

Table 2. The linguistic truth table of *a* OR *b* ($a \vee b$)

a \ *b*	*true*	*nearly true*	*undecided*	*nearly false*	*false*
true	*true*	*true*	*true*	*true*	*true*
nearly true	*true*	*nearly true*	*nearly true*	*nearly true*	*nearly true*
undecided	*true*	*nearly true*	*undecided*	*undecided*	*undecided*
nearly false	*true*	*nearly true*	*undecided*	*nearly false*	*nearly false*
false	*true*	*nearly true*	*undecided*	*nearly false*	*false*

2.3.3 Natural Operator NOT

The third operator NOT is somewhat more involved. In order to achieve our objective, we must define two novel two-tuples set of which the first tuple is linguistic truth variable and the second tuple is its associated ordering index. The first such 2-tuples set is a decreasing 2-tuples set, D_L, which is shown as follows;

$$D_L = \{(l_i, i), i = 1, 2, \ldots, m\}, \tag{7}$$

where i is corresponded order number of l_i in (l_i, i), i = 1, 2, …, m, and $P(l_i) > P(l_{i+1})$, i = 1, 2, …, m – 1.

For example, 2-tuples decreasing set for linguistic truth set L_1 can now be expressed as follows;

$$D_{L_1} = \{(\textit{true}, 1), (\textit{nearly true}, 2), (\textit{undecided}, 3), (\textit{about false}, 4), (\textit{false}, 5)\}.$$

Similarly, we can definite the 2-tuples increasing set G_L accordingly in the following manner;

$$G_L = \{(l_i, i), i = 1, 2, \ldots, m\}, \tag{8}$$

where i is the corresponded order number of l_i in (l_i, i), i = 1, 2, …, m, and $P(l_i) < P(l_{i+1})$, i = 1, 2, …, m – 1.

For example, the increasing linguistic truth set in terms the degree of the truthfulness can be expressed in the following equation. Note that the ordering of the ranking reflects increasing and the degree of truthfulness also is increasing.

$$G_{L_1} = \{(\textit{false}, 1), (\textit{nearly false}, 2), (\textit{undecided}, 3), (\textit{nearly true}, 4), (\textit{true}, 5)\}.$$

An ordering function, ORDER, is also defined for the purpose of finding the correct order of the truth value, say a, as follows:

$$k = \text{ORDER}(a, D_L), \tag{9}$$

where k is order number of *a* in (*a*, k), and $(a, k) \in D_L$.

We also define the linguistic complement function, CMP, with the assistance of a triple valuation equation .The complement linguistic truth of a can then be determined with the help of all three equations (7) ~ (9).

Moreover, linguistic complement function, CMP, in which can get complement linguistic truth value of linguistic truth value, by using both of the above D_L and G_L set, and ORDER function, equations (7) ~ (9) respectively, is defined as

$$\overline{a} = \text{CMP}\ (a, D_L, G_L), \tag{10}$$

where $\overline{a}$ is complement linguistic truth value of a in L, $(\overline{a}, \text{k}) \in G_L$ when $\text{k} = \text{ORDER}\ (a, D_L)$ and $\overline{a} \in L$.

For example, assume $\overline{a}$ is linguistic complement truth value of $a = true$ in L_1, then $\overline{a}$ can be obtained by using equation (10) on the above D_{L_1} and G_{L_1} as

$$\overline{a} = \text{CMP}\ (true, D_{L_1}, G_{L_1}) = false.$$

In the following, NOT linguistic logic operation defined by the above CMP function is introduced. Suppose symbol ~ be denoted to NOT natural operation, then NOT natural operation of a, NOT a or $\sim a$, is expressed as

$$\sim a = \text{CMP}\ (a, D_L, G_L), \tag{11}$$

where a in L.

For a in L_1, we can now complete the linguistic truth table of $\sim a$ as shown in Table 3.

Table 3. The linguistic truth table of NOT a ($\sim a$)

a	$\sim a$
true	*false*
nearly true	*nearly false*
undecided	*undecided*
nearly false	*nearly true*
false	*true*

2.3.4 Natural Operator IF-THEN

In this section, we introduce IF–THEN linguistic logical operation by using the above NOT and OR natural operations. Suppose symbol $\rightarrow$ be presented to IF–THEN natural operation, then IF a THEN b, $a \rightarrow b$, is given by

$$a \rightarrow b = \sim a \vee b. \tag{12}$$

By the above $a \to b$ based on L_1, because of the possible combination of a and b has 16 items, then we only give the simple part of linguistic truth table of $a \to b$ as shown in Table 4.

Table 4. The simple part of linguistic truth table of IF a THEN b ($a \to b$)

a	b	$\sim a$	$a \to b\ (\sim a \vee b)$
true	*true*	*false*	*true*
true	*nearly true*	*false*	*nearly true*
true	*undecided*	*false*	*undecided*
true	*nearly false*	*false*	*nearly false*
true	*false*	*false*	*false*
. .	. .	. .	. .
false	*true*	*true*	*true*
false	*nearly true*	*true*	*true*
false	*undecided*	*true*	*true*
false	*nearly false*	*true*	*true*
false	*false*	*true*	*true*

Now we have established all four tables as show above. Not surprisingly, we are able to find an example which literally moving a novel logically system a whole level upward. Hence we can also claim this is indeed is a closed system, or so called 'grounding' in semiotics jargon.

2.3.5 Properties of Natural Operations

In this section, we are able to establish some interesting properties paralleling those of classical logic. Only some limited and obvious properties of natural logical operations are presented as follows:

Some properties of our above proposed natural operations are discussed as follows.

1) $a = \sim(\sim a)$. (13)

2) $\sim(a \wedge b) = \sim a \vee \sim b$. (14)

3) $\sim(a \vee b) = \sim a \wedge \sim b$. (15)

4) $(a \wedge b) \wedge c = a \wedge (b \wedge c)$. (16)

5) $(a \vee b) \vee c = a \vee (b \vee c)$. (17)

6) $(a \wedge b) \vee c = (a \vee c) \wedge (b \vee c)$. (18)

7) $(a \vee b) \wedge c = (a \wedge c) \vee (b \wedge c)$. (19)

By invoking above 4 tables, it is easy to establish more tables as shown in the following two tables. Obviously, more, much more truth tables for natural logical operations are possible!

First, we prove 1) $a = \sim(\sim a)$ by using the above NOT natural operation, equation (11). The linguistic truth table of $\sim(\sim a)$ is summarized in Table 5.

Table 5. The linguistic truth table of $\sim(\sim a)$

a	$\sim a$	$\sim(\sim a)$
true	*false*	*true*
nearly true	*nearly false*	*nearly true*
undecided	*undecided*	*undecided*
nearly false	*nearly true*	*nearly false*
false	*true*	*false*

Based on the above results in Table 5, we can see that $a = \sim(\sim a)$, that is, property 1) is proved.

Second, property 2) $\sim(a \wedge b) = \sim a \vee \sim b$ is proved by using the above linguistic logical operations of AND, OR, and NOT, equation (5), (6), and (11) respectively. Here, we only give the simple part of linguistic truth table of $\sim(a \wedge b)$ and $\sim a \vee \sim b$. Some results of natural operations are summarized in Table 6.

By the results in Table 6, we can see that $\sim(a \wedge b) = \sim a \vee \sim b$, that is, Property 2) is proved.

Similarly, properties 3) ~ 7) can also be proved by our presented natural operations and its linguistic truth tables.

3 Linguistic Truth Relation by a Linguistic Implication

3.1 Linguistic Logic Propositions

With the development of the linguistic logic's truth tables and their properties, now is the time for us to move on to develop more theory pertinent to potential applications. By doing so, we are in a position to explore the possibility of embarking the important issue of CWW, Computing With Words. The fundamental difference between classical propositions and linguistic logic propositions is in the range of their truth values. While each classical proposition is required to be either true or false, the truth or falsity of linguistic logic proposition is a linguistic truth value.

In this section, we have chosen three simple linguistic logic propositions for closer examination;

1. linguistic logic propositions,
2. conditional linguistic logic propositions,
3. compound linguistic logic propositions.

Table 6. The linguistic truth table of $\sim(a \wedge b)$ and $\sim a \vee \sim b$

a	b	$\sim a$	$\sim b$	$a \wedge b$	$\sim(a \wedge b)$	$\sim a \vee \sim b$
true	*true*	*false*	*false*	*true*	*false*	*false*
true	*nearly true*	*false*	*nearly false*	*nearly true*	*nearly false*	*nearly false*
true	*undecided*	*false*	*undecided*	*undecided*	*undecided*	*undecided*
true	*nearly false*	*false*	*nearly true*	*nearly false*	*nearly true*	*nearly true*
true	*false*	*false*	*true*	*false*	*true*	*true*
.	.	.	.	.	.	.
.	.	.	.	.	.	.
false	*true*	*true*	*false*	*false*	*true*	*true*
false	*nearly true*	*true*	*nearly false*	*false*	*true*	*true*
false	*undecided*	*true*	*undecided*	*false*	*true*	*true*
false	*nearly false*	*true*	*nearly true*	*false*	*true*	*true*
false	*false*	*true*	*true*	*false*	*true*	*true*

The canonical form of linguistic logic propositions of the first type, p_1, is expressed by the statement [9]

$$p_1: \ \aleph \text{ is } F, \tag{20}$$

where $\aleph$ is a variable that takes values n from some universal set N, and F is linguistic set on N that represents a linguistic predicate, such as tall, low, expensive, and so on. Given a particular value $\aleph$ (say, n), this value belongs to F with membership grade $F(n)$. This membership grade is then interpreted as the degree of truth, $T(p_1)$, of linguistic logic proposition p_1. That is,

$$T(p_1) = F(n), \tag{21}$$

for each given particular value n of variable $\aleph$ in linguistic proposition p_1. This means that T is in effect a linguistic truth set, which assigns the membership grade F (n) to each value n of variable $\aleph$.

To illustrate the introduced concepts, let $\aleph$ be the temperature and give a membership function represented, in a given context, the predicate *high*. Then, assuming that all relevant measurement specifications regarding the temperature are given, the corresponding linguistic logic proposition, p_1, is expressed by the statement

$$p_1: \ \textit{temperature} \text{ is } \textit{high}.$$

The degree of truth, $T(p_1)$, depends on the actual value of the temperature and on the given definition of the predicate *high*.

In addition, we consider the second type by the canonical form

$$p_2: \ \text{If } x \text{ is } a, \text{ then } y \text{ is } b, \tag{22}$$

where x, y are linguistic logic propositions of the above first type, whose values are in L. The degree of truth, $T(p_2)$, can be considered by using the IF–THEN linguistic logic operation in equation (12) as

$$T(p_2) = \sim a \vee b, \tag{23}$$

where a, b are the degrees of truth of x, y, respectively, and a, b, and $T(p_2) \in L$.

For example, let p_2 be a conditional linguistic logic proposition as

p_2: If *temperature* is *high* is *about true*, then *weather* is *hot* is *true*,

then linguistic truth value of p_2 in L_1, $T(p_2)$, is *true* by using natural operations of NOT (~) and OR ($\vee$) in equation (23).

Furthermore, assume q and r are linguistic logic propositions or conditional linguistic logic propositions, the compound linguistic logic proposition of the third type, p_3, is discussed by the canonical form as following.

$$p_3\text{:}\;\; q * r, \tag{24}$$

where $*$ denotes a linguistic logic operation, for example, AND (equation (5)), OR (equation (6)), and so on. The degree of truth, $T(p_3)$, can be defined as

$$T(p_3) = T(q) * T(r),$$

where $T(q)$, $T(r)$ are the degree of truth of q and r, respectively, and $T(q)$, $T(r)$, and $T(p_3) \in L$.

For example, let q, r be two linguistic logic propositions as

q: *tomato* is *small*,

r: *orange* is *large*.

Then, the some examples of compound linguistic logic propositions by using the above q and r can be expressed by

p_{3_1}: *tomato* is *small* AND *orange* is *large*,

p_{3_2}: *tomato* is *small* OR *orange* is *large*,

3.2 Linguistic Implication

The logic operation of implication is the back bone for approximate reasoning for linguistic logic implication just like that of the classical logic approximated reasoning. In general, linguistic implication, η, is a function of the form

$$\eta : L \times L \to L,$$

which for any possible truth values s, t of given linguistic logic propositions x, y, respectively, defines the truth value, $\eta(s, t)$, of the conditional linguistic proposition "if x, then y." This function should be an extension of the classical implication, $x \rightarrow y$, from the restricted domain {0, 1} to the full domain [0, 1] of linguistic truth values in linguistic logic.

In classical logic [9], where $s, t \in \{0, 1\}$, η can be defined in several distinct forms. While these forms are equivalent in classical logic, their extensions to linguistic logic are not equivalent and result in distinct classes of linguistic implications. This fact makes the concept of linguistic implication somewhat complicated.

One method to evaluate the implication operation, in classical logic implication, is through the operations of its equivalence such as follows;

$$\eta\,(s, t) = \bar{s} \vee t,$$

where $\eta\,(s, t)$ is the truth value of the classical conditional proposition "If s, then t", for all $s, t \in \{0, 1\}$. This function should be an extension of the classical implication, $s \rightarrow t$, from the restricted domain {0, 1} to the full domain [0, 1] of linguistic truth values in linguistic logic.

We discover the best way to generalize the implication in linguistic logic implication is to use the same equivalence. In this paper, we interpret the disjunction and negation as a linguistic logic union and linguistic logic complement, respectively. This results in defining $\eta\,(a, b)$ in linguistic logic by the equation

$$\eta\,(a, b) = \sim a \vee b, \tag{25}$$

where a, b are variables whose values are in L, and $\vee$ and $\sim$ denote OR natural operation in equation (6) and NOT natural operation in equation (11), respectively.

For example, assume that $\eta\,(a, b) = \sim a \vee b$ be a linguistic implication defined on $L_1 \times L_1$ in which a and b are in L_1. Then, the results of $\eta\,(a, b)$ are shown in Table 7.

Table 7. The results of $\eta\,(a, b) = \sim a \vee b$

a \ b	*true*	*nearly true*	*undecided*	*nearly false*	*false*
true	*true*	*nearly true*	*undecided*	*nearly false*	*false*
nearly true	*true*	*nearly true*	*undecided*	*nearly false*	*nearly false*
undecided	*true*	*nearly true*	*undecided*	*undecided*	*undecided*
nearly false	*true*	*nearly true*	*nearly true*	*nearly true*	*nearly true*
false	*true*	*true*	*true*	*true*	*true*

The above linguistic implication is obtained by generalized the implication operator of classical logic as mentioned above. It is quite amazing that Table 7 reduces to the same entries as that of the classical logic implications should we limited a and b to only truth values to only to true and false! Identifying various properties of the classical implication and generalizing them appropriately leads to the following properties, which are viewed as reasonable axioms of our above linguistic implication, η (a, b).

Axiom 1. $a_1 \leq b_2$ implies η $(a_1, b) \geq$ η (a_2, b) (monotonicity in first argument).
Axiom 2. $b_1 \leq b_2$ implies η $(a, b_1) \geq$ η (a, b_2) (monotonicity in second argument).
Axiom 3. η (*false*, b) = *true* (dominance of falsity).
Axiom 4. η (*true*, b) = b (neutrality of truth).
Axiom 5. η $(a_1$, η $(a_2, b)) \geq$ η $(a_2$, η $(a_1, b))$ (exchange property).
Axiom 6. η is a continuous function (continuity).

3.3 Linguistic Truth Relation

A crisp relation represents the presence or absence of association, interaction or interconnectedness between the elements of two or more sets. The concept can be generalized to allow for various degree or strengths of association or interaction between elements. Degrees of association can be represented by membership grades in a fuzzy relation in the same way as degrees of set membership are represented in the fuzzy set.

In fact, just as the crisp set can be viewed as a restricted case of the more general fuzzy set concept, the crisp relation can be considered to be a restricted case of the fuzzy relation. Fuzzy relations are fuzzy subsets of X × Y, that is, mappings from X → Y. It has been studied by a number of authors, in particular by Zadeh, Kaufmann, and Rosenfeld. Applications of fuzzy relations are widespread and important.

Binary relations have a special significance among n-dimensional relations since they are, in some sense, generalized mathematical functions. Contrary to function from X to Y, binary relation R(X, Y) may assign to each element of X two or more elements of Y.

In addition to defining a binary fuzzy relation that exists between two different fuzzy sets, it is also possible to define a fuzzy binary relation among the elements of a single fuzzy set X. A binary fuzzy relation of this type can be denoted by R(X, X) or $R(X^2)$ and is a subset of $X \times X = X^2$.

Ordinary fuzzy binary relations (with the valuation set [0, 1]) can obviously be extended to binary linguistic logic relations (with a linguistic truth set) in the same way as fuzzy sets are extended to linguistic truth sets.

In the following, the binary linguistic truth relation defined on a linguistic implication will be discussed on linguistic condition propositions. Let LTR (a, b), η (a, b) be linguistic truth relation and linguistic implication defined on $L \times L$ where a, b are variables whose values are in L, respectively.

Then the linguistic truth relation based on the linguistic implication can be determined for $a, b \in L$ by the equation

$$\text{LTR}\ (a, b) = \eta\ (a, b). \tag{26}$$

From equation (26) and the results of the above η (a, b) in Table 7, we can see that LTR is irreflexive, asymmetric, and nontransitive.

3.4 Linguistic Approximated Reasoning [LAR]

It is important to recall our original assumption at very outset to be a very limited set of words for our language. Further assumption by choosing the fuzzy sets judiciously allow this language to be a close set in the sense there are five words exactly. Hence, we have a well grounded semiotics system with only finite words. Also, we are able to construct a well behaved linguistic logic system at the first place. Natural language in general, in the sense of Noam Chomsky's theory, the words can be infinitive and with a grammar. CWW or Fuzzy Logic with limited words already finds realization or practical usage in industry [11] at present. Hence this very limited linguistic logic system with only five words may play a very useful role as a model for truthfulness may have its niche in application in the real world.

In this section, we will restrict our consideration to generalized modus ponens and generalized modus tollens. For a generalized modus ponens case, we first introduce linguistic approximation reasoning under conditional linguistic logic propositions by using the above linguistic truth relation, equation (26), and some natural operations of linguistic logic.

For example, let x, y be linguistic logic propositions, and a, b, a', b' be linguistic truth values in L. Then the generalized modus ponens can be expressed by

Premise	x is a'
Implication	If x is a, then y is b
Conclusion	y is b'

Assume that conditional linguistic proposition p of the form is given by

p: If x is a , then y is b,

where x, y be linguistic logic propositions, a and b be the linguistic truth values of x and y in L, respectively. Let the form "x is a'," be a premise, and we then want to make a conclusion in the form "y is b'," for $a', b' \in L$.

Assume that LTR is a linguistic truth relation defined on $L \times L$, and determined by the linguistic implication in equation (26), then if LTR and a' are given, we can obtain b' by compositional rule of linguistic inference as

$$b' = \mathrm{MIN}\,(a', \mathrm{LTR}\,(a, b)), \tag{27}$$

where MIN is a minimum function in equation (3).

In this linguistic inference, we can see that the generalized modus ponens becomes the classical modus ponens when the linguistic truth set is crisp, $L = \{$*true*, *false*$\}$.

For example, assume that conditional linguistic proposition "If x is a, then y is b" is given, where a = *true* and b = *true* in L_1. Then, given a premise expressed by the linguistic proposition "x is a'," where a' = *nearly true*, we want to use the generalized modus ponens to derive a conclusion in the form "y is b'."

By equation (26) and the results of the linguistic implication, η (a, b) in the above Table 7, the linguistic truth relation LTR defined on $L_1 \times L_1$ can be obtained. Then, b' can be obtained by the compositional rule of linguistic inference in equation (27),

$$\begin{aligned} b' &= \text{MIN}\ (\textit{nearly true}, \text{LTR}\ (\textit{true}, \textit{true})) \\ &= \text{MIN}\ (\textit{nearly true}, \textit{true}) \\ &= \textit{nearly true}. \end{aligned}$$

In addition, another linguistic inference rule, which is a generalized modus tollens, is expressed by the form

Premise	y is b'
Implication	If x is a, then y is b
Conclusion	x is a'

Then, compositional rule of linguistic inference of this type can be given

$$a' = \text{MIN}\ (b', \text{LTR}\ (a, b)). \tag{28}$$

In this linguistic inference, we can see that the generalized modus ponens becomes the classical modus ponens when the linguistic truth set is crisp, L = {*true*, *false*}.

For example, assume that conditional linguistic proposition "If x is a, then y is b" is given, where a = *true* and b = *true* in L_1. Then, given a premise expressed by the linguistic proposition "y is b'," where b' = *false*, we want to use the generalized modus ponens to derive a conclusion in the form "x is a'."

By the same above way, a' can be obtained by the compositional rule of linguistic inference in equation (28),

$$\begin{aligned} a' &= \text{MIN}\ (\textit{false}, \text{LTR}\ (a, b)) \\ &= \text{MIN}\ (\textit{false}, \textit{true}) \\ &= \textit{false}. \end{aligned}$$

Once again, we would like to stress the reason the approximated reasoning can be carried out neatly and orderly in this paper is due to our restricted and closed system of a set of finite words of 5 in our dictionary. The approximated reasoning in the wide sense necessarily is expected to deal to many uncertainty situation of which the system is expected to handle new words and new situation. Natural language inevitably must be evolutionary! So called CWW, in general is expected to perform difficulty situation under uncertainty environment to be robust. As one can surmise, it is a long way to achieve these noble goals.

4 Concluding Remarks

The work presented in this paper has been motivated primarily by Zadeh's idea of linguistic variables intended to provide rigorous mathematical modeling of natural language, approximated reasoning and CWW, Computing With Words [15-16, 18-19]. We must understand the magnitude of influence should these idea can indeed be successfully being carried out with huge potential applications in biological science, social science and the like [16, 19]. The potential of possible applications of these theory, especially the approximated reasoning in decision making, which in term in the areas such as the managements, operations research and system engineering [1, 15] indeed can be immensely important!

The immediately goals of this paper have been two folds: (i) The modeling of truthfulness which is badly needed in pragmatic applications. (ii) To test the idea of semiotics by moving the handling of a chunk of information to a higher level by way of fuzzy set concept. We feel we have achieved our objective with a somewhat limited assumed model with only five words. In the meantime, some already published works, [2, 4, 5], have facilitated our further applications of their results. Specifically, we have accomplished the following three subgoals: (i) Some natural operations based on computing with linguistic truth values are used to process linguistic approximate reasoning and linguistic logic. (ii) The linguistic truth relation determined by the linguistic implication whose some axioms are viewed as reasonable properties is irreflexive, asymmetric, and nontransitive. (iii) The simple linguistic compositional rules applied to the linguistic inferences of generalized modus ponens and generalized modus tollens on conditional linguistic propositions, equation (27) and (28) respectively, are defined on the linguistic truth relation and some natural operations.

CWW is a huge and ambitious goal which obviously will take forever long period of time to reach a satisfactory results because it inevitably will involve the same struggle as those natural language and linguistic researchers have had done over the centuries long endeavor! Semantics issues alone would a huge obstacle to overcome. The novel challenge here is to have CWW operate like a specialized computer with natural language inputs and outputs.

The realistic immediate future research would be the linguistic inference of generalized hypothetical syllogism which is an interesting research topic to explore. In that capacity, linguistic logic ought to be studied via the compound linguistic truth relations. After we have accomplished these immediate goals, then the power of CWW as applied to business decision making, management with humanistic flavor artificial intelligence, even the natural language applications to social sciences such as law, income tax return, etc. can be seen. When that day come, the significance of the urging in further development of the mathematics of uncertainty will be transparently clear.

References

1. Carlsson, C., Fedrizzi, M., Fuller, R.: Fuzzy logic in management, p. 41. Kluwer Academic Publishers, Dordrecht (2004)
2. Chen, S.H., Hsieh, C.H.: Graded mean integration representation of generalized fuzzy number. In: Proceedings of the sixth Conference of Chinese Fuzzy Sets and Systems (CD-ROM: filename: 031.wdl), p. 31 (1998)
3. Chen, S.H., Hsieh, C.H.: Ranking generalized fuzzy number with graded mean integration representation. In: Proceedings of the Eighth International Conference of Fuzzy Sets and Systems Association World Congress, vol. 2, pp. 551–555 (1999)
4. Chen, S.H., Hsieh, C.H.: Representation, Ranking, Distance, and Similarity of L-R type Fuzzy Number and Application. Australian Journal of Intelligent Information Processing Systems 6(4), 217–229 (2000)
5. Hsieh, C.H., Chen, S.H.: Similarity of generalized fuzzy number with graded mean integration representation. In: Proceedings of the Eighth International Conference of Fuzzy Sets and Systems Association World Congress, vol. 2, pp. 899–902 (1999)
6. Hsieh, C.H.: The natural operations of linguistic logic. New Mathematics and Natural Computations 4(1), 77–86 (2008)
7. Hintikka, J., Suppes, P.: Aspects of inductive logic. North-Holland Publ. Co., Amsterdam (1996)
8. Kaufmann, A., Gupta, M.M.: Introduction to fuzzy arithmetic theory and applications. Int. Thomson Computer Press, USA (1991)
9. Klir, G.J., Yuan, B.: Fuzzy sets and fuzzy logic—Theory and applications. Prentice-Hall Inc., Englewood Cliffs (1995)
10. Łukasiewicz, J.: Aristotle's syllogistic. Clarendon Press, Oxford (1951)
11. Lawry, J., G. Shanahan, J., L. Ralescu, A. (eds.): Modelling with Words. LNCS (LNAI), vol. 2873. Springer, Heidelberg (2003)
12. Rescher, N.: Many-valued logic. McGraw-Hill, New York (1969)
13. Wang, P.P. (ed.): Advances in Fuzzy Theory and Technology (USA) III, 3 (1997)
14. Wang, P.P.: Computing with Words. John Wiley & Sons, Inc., Chichester (2001)
15. Zadeh, L.A.: Fuzzy logic-Computing with word. IEEE Trans. Fuzzy Syst. 4, 103–111 (1996)
16. Zadeh, L.A.: Fuzzy logic and approximate reasoning. In: Klir, G.J., Yuan, B. (eds.) Fuzzy sets and fuzzy logic—Theory and applications, World Scientific Publishing Co. Pte. Ltd., Singapore (1996)
17. Zadeh, L.A.: What is truth? BISC Group Seminar & Discussion, BISC, University of California at Berkeley. Berkeley Initiative in Soft Computing, June 2 (2009)
18. Zimmermann, H.J.: Fuzzy set theory and its applications, 2nd edn. Kluwer Academic Publishers, Dordrecht (1991)
19. Ching, H.: Translated by Sister Mary Lelia Makra. St. John's University Press, New York (1961)

A Nonlinear Integral Which Generalizes Both the Choquet and the Sugeno Integral*

Erich Peter Klement[1], Radko Mesiar[2,3], and Endre Pap[4]

[1] Department of Knowledge-Based Mathematical Systems, Johannes Kepler University, 4040 Linz, Austria
e-mail: ep.klement@jku.at

[2] Department of Mathematics and Descriptive Geometry, Faculty of Civil Engineering, Slovak University of Technology, 81 368 Bratislava, Slovakia
e-mail: mesiar@math.sk

[3] Institute for Research and Applications of Fuzzy Modeling, University of Ostrava, 70103 Ostrava, Czech Republic

[4] Department of Mathematics and Informatics, University of Novi Sad, 21000 Novi Sad, Serbia
e-mail: pap@dmi.uns.ac.rs, pape@eunet.rs

Abstract. The Choquet and the Sugeno integral provide a useful tool in many problems in engineering and social choice where the aggregation of data is required. However, their applicability is somehow restricted because of the special operations used in the construction of these integrals. This survey presents the main ideas and results concerning the construction of a universal integral generalizing both the Choquet and the Sugeno case. For functions with values in the nonnegative real numbers, universal integrals can be defined on arbitrary measurable spaces and for arbitrary monotone measures. A more detailed exposition and the proofs of the propositions can be found in [38].

Keywords: Universal integral, Choquet integral, Sugeno integral, aggregation, pseudo-multiplication.

1 Motivation

Two early concepts for measuring fuzzy sets [74] by means of (different) integrals were presented by L. A. Zadeh [75] and M. Sugeno [66]. In [75], the starting point was a probability measure P on a σ-algebra of subsets of the universe X. For a fuzzy event F, i.e., a fuzzy subset F of X with measurable membership function μ_F

* The second and the third author were supported by a technology transfer project of the Upper Austrian Government and by project SK-SRB-19. The second author was also supported by the grants APVV-0375-06 and MSM VZ 6198898701. The third author was also supported by Vojvodina PSSTD, by the grants MNTRS-144012 and MTA of HTMT, and by the French-Serbian project "Pavle Savić".

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 39–52.
springerlink.com

(*vertical representation* [50] of F), the *probability* $\tilde{P}(F)$ was defined by the standard Lebesgue integral

$$\tilde{P}(F) = \int_X \mu_F \, dP. \tag{1}$$

This concept, which has been investigated extensively and generalized in [5, 6, 7, 30, 32, 34], has some limitations because of the (σ-)additivity of the underlying probability measure P.

In [66], the *horizontal representation* [50] by means of the *level sets* $(F_t)_{t\in[0,1]}$ of a fuzzy subset F of X, where

$$F_t = \{x \in X \mid \mu_F(x) \geq t\}, \tag{2}$$

and the *Sugeno integral* based on some continuous monotone set function (see (4) below) were used to measure the fuzzy set F. It was shown (see, e.g., [4, 11, 19, 23, 49, 53, 55]) that this approach can be extended to rather general monotone set functions.

The situation is similar for the *Choquet integral* (see (3) below) [4, 10, 19, 62, 69, 70] which is based on not necessarily additive measures (*capacities*) and on level sets of the integrand. Moreover, the Choquet integral covers also the classical Lebesgue integral. As a consequence, the probability of a fuzzy event (1) can also be obtained using a Choquet integral. A horizontal approach to integration can also be found in [40, Chapter 4], leading again to the Choquet integral.

For both the Choquet and the Sugeno integral, the horizontal representation (2) of a fuzzy subset F of X can be readily extended to arbitrary non-negative measurable functions. Moreover, this horizontal approach gives some freedom in the choice of the character of the underlying measure. On the other hand, the Choquet integral is based on the standard arithmetic operations addition and multiplication, implying the comonotone additivity for each Choquet integral-based aggregation. Since the Sugeno integral is based on the lattice operations join and meet, the corresponding aggregation is necessarily join- and meet-homogeneous. However, not all aggregation processes in engineering and social choice are comonotone additive or join- or meet-homogeneous. If such aggregations should be done by some integral, other types of integral different from the Choquet and the Sugeno integral should be at hand.

Here we survey the concept of universal integrals acting on the interval $[0,\infty]$ and generalizing both the Choquet and the Sugeno integral, which were inspired by the horizontal approach to integration (for more details see [38]). These integrals can be defined on an arbitrary measurable space $(X,\mathcal{A})$, they are based on an arbitrary monotone measure $m\colon \mathcal{A} \to [0,\infty]$, and they can be applied to arbitrary measurable functions $f\colon X \to [0,\infty]$ using their level sets.

For us, a *monotone measure* m on a measurable space $(X,\mathcal{A})$ is a function $m\colon \mathcal{A} \to [0,\infty]$ satisfying $m(\emptyset) = 0$, $m(X) > 0$, and $m(A) \leq m(B)$ whenever $A \subseteq B$. Note that a monotone measure is not necessarily (σ-)additive. This concept goes back to M. Sugeno [66] (where also the continuity of the measures was required). To be precise, normed monotone measures on $(X,\mathcal{A})$, i.e., monotone measures satisfying $m(X) = 1$, are also called *fuzzy measures* [19, 66, 70, 71] or *capacities* [19] or (*monotone*) *games* [2, 7], depending on the context. Measurable functions from X to $[0,1]$ can be considered as (membership functions of) *fuzzy events* in $(X,\mathcal{A})$ [29, 30, 31, 66, 70, 75]. Accordingly, Sugeno and Choquet integrals on the scale $[0,1]$ are often called *fuzzy integrals* [19, 47] (see also [37]). Extensions of these integrals to functions whose range is a bipolar scale (e.g., the interval $[-1,1]$ or the whole real line $\mathbb{R}$) exist (see, e.g., [16]).

Sugeno and Choquet integrals proved to be useful in several applications (for an overview see [17]), notably in some problems of game theory and multicriteria decision making, where the score vectors describing the single alternatives are modeled by fuzzy subsets of the set of all criteria, and the monotone measures quantify the weight of sets of criteria [13, 15, 19, 22, 57, 58, 73]. Other applications concern aggregation operators [9, 42, 43] as well as pattern recognition and classification [20, 28] and information theory [41]. A particularly popular example of a Sugeno integral is the so-called Hirsch index (h-index) [68] which measures the cumulative impact of a researcher's output by looking at the amount of citation his or her work has received.

The fact that both the Sugeno and the Choquet integral are in a close relationship to aggregation functions is well-known. To give a recent example, in [21] the authors write about the Sugeno and the Choquet integral:

> "Their mathematical properties as aggregation functions have been studied extensively [...], and it is known that many classical aggregation functions are particular cases of these so-called fuzzy integrals, e.g., the weighted arithmetic mean, ordered weighted averages (OWA), weighted minimum and maximum, etc."

2 Universal Integrals

For a fixed measurable space $(X,\mathcal{A})$, i.e., a non-empty set X equipped with a σ-algebra $\mathcal{A}$, recall that a function $f\colon X \to [0,\infty]$ is called $\mathcal{A}$-measurable if, for each $B \in \mathcal{B}([0,\infty])$, the σ-algebra of Borel subsets of $[0,\infty]$, the preimage $f^{-1}(B)$ is an element of $\mathcal{A}$. We shall use the following notions:

Definition 1. *Let $(X,\mathcal{A})$ be a measurable space.*

(i) $\mathcal{F}^{(X,\mathcal{A})}$ *denotes the set of all $\mathcal{A}$-measurable functions $f\colon X \to [0,\infty]$;*
(ii) for each number $a \in\,]0,\infty]$, $\mathcal{M}_a^{(X,\mathcal{A})}$ *denotes the set of all monotone measures satisfying $m(X) = a$, and we put*

$$\mathcal{M}^{(X,\mathcal{A})} = \bigcup_{a\in]0,\infty]} \mathcal{M}_a^{(X,\mathcal{A})}.$$

Each non-decreasing function $H\colon \mathcal{F}^{(X,\mathcal{A})} \to [0,\infty]$ with $H(\mathbf{0}) = 0$ is called an *aggregation function on* $\mathcal{F}^{(X,\mathcal{A})}$ (compare [9, 18], in particular [8]). Which aggregation function "deserves" to be called an integral, this is a classical and still controversial question. We give three examples of such functions which are known and used as integrals.

The *Choquet* [10], *Sugeno* [66] and *Shilkret* [61] integrals (see also [4,54]), respectively, are given, for any measurable space $(X,\mathcal{A})$, for any measurable function $f \in \mathcal{F}^{(X,\mathcal{A})}$ and for any monotone measure $m \in \mathcal{M}^{(X,\mathcal{A})}$, by

$$\mathbf{Ch}(m,f) = \int_0^\infty m(\{f \geq t\})\,dt, \tag{3}$$

$$\mathbf{Su}(m,f) = \sup\{\min(t, m(\{f \geq t\})) \mid t \in [0,\infty]\}, \tag{4}$$

$$\mathbf{Sh}(m,f) = \sup\{t \cdot m(\{f \geq t\}) \mid t \in [0,\infty]\}, \tag{5}$$

where the convention $0 \cdot \infty = 0$ is adopted whenever necessary.

Independently of whatever measurable space $(X,\mathcal{A})$ is actually chosen, all these integrals map $\mathcal{M}^{(X,\mathcal{A})} \times \mathcal{F}^{(X,\mathcal{A})}$ into $[0,\infty]$ and, fixing an arbitrary $m \in \mathcal{M}^{(X,\mathcal{A})}$, they are aggregation functions on $\mathcal{F}^{(X,\mathcal{A})}$. Moreover, fixing an arbitrary $f \in \mathcal{F}^{(X,\mathcal{A})}$, they are non-decreasing functions from $\mathcal{M}^{(X,\mathcal{A})}$ into $[0,\infty]$.

Let $\mathcal{S}$ be the class of all measurable spaces, and put

$$\mathcal{D}_{[0,\infty]} = \bigcup_{(X,\mathcal{A})\in\mathcal{S}} \mathcal{M}^{(X,\mathcal{A})} \times \mathcal{F}^{(X,\mathcal{A})}.$$

Each of the integrals mentioned in (3)–(5) maps $\mathcal{D}_{[0,\infty]}$ into $[0,\infty]$ and is non-decreasing in each coordinate.

For the definition of our integral we shall need a pseudo-multiplication (mentioned first in [67]). Note that pseudo-multiplications, in general, are neither associative nor commutative.

A function $\otimes\colon\ [0,\infty]^2 \to [0,\infty]$ is called a *pseudo-multiplication* if it satisfies the following properties:

(i) $\otimes$ is non-decreasing in each component, i.e., for all $a_1,a_2,b_1,b_2 \in [0,\infty]$ with $a_1 \leq a_2$ and $b_1 \leq b_2$ we have $a_1 \otimes b_1 \leq a_2 \otimes b_2$;
(ii) 0 is an annihilator of $\otimes$, i.e., for all $a \in [0,\infty]$ we have $a \otimes 0 = 0 \otimes a = 0$;
(iii) $\otimes$ has a neutral element different from 0, i.e., there exists an $e \in \,]0,\infty]$ such that, for all $a \in [0,\infty]$, we have $a \otimes e = e \otimes a = a$.

All three integrals mentioned in (3)–(5) fulfill the equality $\mathbf{I}(m_1, f_1) = \mathbf{I}(m_2, f_2)$ whenever the pairs $(m_1, f_1), (m_2, f_2) \in \mathcal{D}_{[0,\infty]}$ satisfy, for all $t \in \,]0,\infty]$,

$$m_1(\{f_1 \geq t\}) = m_2(\{f_2 \geq t\}). \tag{6}$$

Therefore, the following equivalence relation on the class $\mathcal{D}_{[0,\infty]}$ makes sense in our context: two pairs $(m_1, f_1) \in \mathcal{M}^{(X_1,\mathcal{A}_1)} \times \mathcal{F}^{(X_1,\mathcal{A}_1)}$ and $(m_2, f_2) \in \mathcal{M}^{(X_2,\mathcal{A}_2)} \times \mathcal{F}^{(X_2,\mathcal{A}_2)}$ are called *integral equivalent* (in symbols $(m_1, f_1) \sim (m_2, f_2)$) if they satisfy property (6).

The integrals given in (3)–(5) do not distinguish between integral equivalent pairs. This observation has motivated us to take the indistinguishability of integral equivalent pairs as an axiom for our integrals.

We also require universality of the integral, in the sense that it can be defined on any measurable space $(X,\mathcal{A})$. Therefore we will use the name universal integral in what follows. Summarizing, this led to the following axiomatic approach for universal integrals:

Definition 2. *A function* $\mathbf{I}\colon \mathcal{D}_{[0,\infty]} \to [0,\infty]$ *is called a* universal integral *if the following axioms hold:*

(I1) *For any measurable space* $(X,\mathcal{A})$*, the restriction of the function* $\mathbf{I}$ *to* $\mathcal{M}^{(X,\mathcal{A})} \times \mathcal{F}^{(X,\mathcal{A})}$ *is non-decreasing in each coordinate;*

(I2) *there exists a pseudo-multiplication* $\otimes\colon [0,\infty]^2 \to [0,\infty]$ *such that for all* $(m, c\cdot \mathbf{1}_A) \in \mathcal{D}_{[0,\infty]}$

$$\mathbf{I}(m, c\cdot\mathbf{1}_A) = c \otimes m(A);$$

(I3) *for all integral equivalent pairs* (m_1, f_1), $(m_2, f_2) \in \mathcal{D}_{[0,\infty]}$ *we have*

$$\mathbf{I}(m_1, f_1) = \mathbf{I}(m_2, f_2).$$

The following observations are immediate consequences of Definition 2. In particular, axioms (I1)–(I3) have rather natural interpretations.

(i) All three integrals mentioned in (3)–(5) are universal integrals in the sense of Definition 2; the underlying pseudo-multiplication $\otimes$ is the standard product (with neutral element 1) in the case of the Choquet and the Shilkret integral, while $\otimes$ is the minimum (with neutral element ∞) for the Sugeno integral.

(ii) Axiom (I1) simply requires the integral not to decrease if the integrand and/or the underlying monotone measure are replaced by a greater one.

(iii) Axiom (I2) is a kind of "truth functionality" (as it is known in propositional logic) adopted to the case of integrals: it means that for a function $c\cdot\mathbf{1}_A\colon X \to [0,\infty]$ (where $\mathbf{1}_A(x) = 1$ if $x \in A$, and $\mathbf{1}_A(x) = 0$ otherwise), the integral of $c\cdot\mathbf{1}_A$ only depends on the two numbers c and $m(A)$ and *not* on the underlying space $(X,\mathcal{A})$.

The fact that e is a left neutral element of $\otimes$ allows us to reconstruct the underlying monotone measure m from $\mathbf{I}$ via $m(A) = \mathbf{I}(m, e\cdot\mathbf{1}_A)$. That e is also a right neutral element of $\otimes$ implies that $\mathbf{I}$ is idempotent in the sense that $\mathbf{I}(m, c\cdot\mathbf{1}_X) = c$, regardless of the measurable space $(X,\mathcal{A}) \in \mathcal{S}$ and the monotone measure m under consideration.

(iv) Finally, axiom (I3) guarantees that the integral does not distinguish between integral equivalent pairs.

(v) Due to axiom (I3), for each universal integral $\mathbf{I}$ and for each pair $(m,f) \in \mathcal{D}_{[0,\infty]}$, the value $\mathbf{I}(m,f)$ depends only on the function $h^{(m,f)}\colon\ ]0,\infty] \to [0,\infty]$ given by

$$h^{(m,f)}(x) = m(\{f \geq x\}).$$

Note that, for each $(m,f) \in \mathcal{D}_{[0,\infty]}$, the function $h^{(m,f)}$ is non-increasing and thus Borel measurable.

Denote by $\mathcal{H}$ the subset of all non-increasing functions from $\mathcal{F}^{(]0,\infty],\mathcal{B}(]0,\infty])}$. The proof of the following characterization can be found in [38].

Proposition 1. *A function* $\mathbf{I}\colon \mathcal{D}_{[0,\infty]} \to [0,\infty]$ *is a universal integral related to some pseudo-multiplication* $\otimes$ *if and only if there is a function* $J\colon \mathcal{H} \to [0,\infty]$ *satisfying the following conditions:*

(J1) *J is non-decreasing;*
(J2) $J(d\cdot \mathbf{1}_{]0,c]}) = c\otimes d$ *for all* $c,d \in [0,\infty]$;
(J3) $\mathbf{I}(m,f) = J(h^{(m,f)})$ *for all* $(m,f) \in \mathcal{D}_{[0,\infty]}$.

An approach to universal integrals similar to Proposition 1 can be traced back to [64], compare also [36]. The following is an abstract example of a universal integral (which is neither of the Choquet nor Sugeno nor Shilkret type) and illustrates the interrelationship between the functions $\mathbf{I}$, $\otimes$, and J: let $\mathbf{I}\colon \mathcal{D}_{[0,\infty]} \to [0,\infty]$ be given by

$$\mathbf{I}(m,f) = \sup\left\{ \tfrac{t\cdot m(\{f\geq t\})}{t+m(\{f\geq t\})} \,\middle|\, t\in\,]0,\infty] \right\}.$$

Then $\mathbf{I}$ is a universal integral. Moreover, we have

$$\mathbf{I}(m, c\cdot \mathbf{1}_A) = \tfrac{c\cdot m(A)}{c+m(A)},$$

i.e., the underlying pseudo-multiplication $\otimes\colon [0,\infty]^2 \to [0,\infty]$ (with neutral element ∞) is given by $a\otimes b = \frac{a\cdot b}{a+b}$, and the function $J\colon \mathcal{H} \to [0,\infty]$ by

$$J(h) = \sup\left\{ \tfrac{t\cdot h(t)}{t+h(t)} \,\middle|\, t\in\,]0,\infty] \right\}.$$

Note that this universal integral is neither comonotone additive nor join- nor meet-homogeneous.

All considerations above can be isomorphically transformed, replacing the interval $[0,\infty]$ (which was chosen here because of its generality) by an interval $[0,b]$ with $b\in\,]0,\infty[$ (compare the approach described in [4]). When restricting our consideration to the interval $[0,1]$ (i.e., $b=1$), an interesting model for subjective evaluation proposed in [26] is based on a kind of integral which covers both the Choquet and the Sugeno integral. This integral is known as *Imaoka integral* and it fits into our framework of universal integrals as given in Definition 2, and the corresponding pseudo-multiplication $\otimes$ is a member of the Frank family of t-norms [14].

3 A Construction Method

For a given pseudo-multiplication $\otimes$ on $[0,\infty]$, we suppose the existence of a pseudo-addition $\oplus\colon [0,\infty]^2 \to [0,\infty]$ which is continuous, associative, non-decreasing and has 0 as neutral element (then the commutativity of $\oplus$ follows, see [35]), and which is left-distributive with respect to $\otimes$, i.e., for all $a,b,c\in[0,\infty]$ we have $(a\oplus b)\otimes c = (a\otimes c)\oplus(b\otimes c)$. The pair $(\oplus,\otimes)$ is then called an *integral operation pair*.

Note that similar operations for the construction of an integral were considered in [25,52,67], but for the vertical representation of functions. Therefore, these integrals are limited to $\oplus$-additive measures only, and they do not fit into our framework of universal integrals.

For each measurable space $(X,\mathcal{A})$ and for each *simple* function $s \in \mathcal{F}^{(X,\mathcal{A})}$, i.e., s has finite range $\mathrm{Ran}(s) = \{a_1, a_2, \dots, a_n\}$ where $a_1 < a_2 < \cdots < a_n$, using the notations $A_i = \{s \geq a_i\}$, $a_0 = 0$, and $b_i = \inf\{c \in [0,\infty] \mid a_{i-1} \oplus c = a_i\}$ we have the following unique representation:

$$s = \bigoplus_{i=1}^{n} b_i \cdot \mathbf{1}_{A_i}. \tag{7}$$

We denote the set of all simple functions in $\mathcal{F}^{(X,\mathcal{A})}$ by $\mathcal{F}^{(X,\mathcal{A})}_{\text{simple}}$.

Using representation (7) for a function $s \in \mathcal{F}^{(X,\mathcal{A})}_{\text{simple}}$, note that the continuity of $\oplus$ implies that $a_{i-1} \oplus b_i = a_i$ for each $i \in \{1,\dots,n\}$. By induction and the associativity of $\oplus$ we get, for each $j \in \{1,\dots,n-1\}$ and $r \in \{1,\dots,n-j\}$, $a_{j-1} \oplus b_j \oplus \cdots \oplus b_{j+r} = a_{j+r}$, and the monotonicity and the continuity of $\oplus$ imply

$$b_j \oplus \cdots \oplus b_{j+r} = \inf\{c \in [0,\infty] \mid a_{j-1} \oplus c = a_{j+r}\}.$$

In analogy to the Lebesgue integral, for each measurable space $(X,\mathcal{A})$ and for each monotone measure m on $(X,\mathcal{A})$ let $\mathbf{I}^{\text{simple}}_{\oplus,\otimes} \colon \mathcal{M}^{(X,\mathcal{A})} \times \mathcal{F}^{(X,\mathcal{A})}_{\text{simple}} \to [0,\infty]$ be the function given by

$$\mathbf{I}^{\text{simple}}_{\oplus,\otimes}\left(m, \bigoplus_{i=1}^{n} b_i \cdot \mathbf{1}_{A_i}\right) = \bigoplus_{i=1}^{n} b_i \otimes m(A_i). \tag{8}$$

Note that for a simple function $s \in \mathcal{F}^{(X,\mathcal{A})}_{\text{simple}}$ there can be several representations

$$s = \bigoplus_{j=1}^{k} c_j \cdot \mathbf{1}_{C_j},$$

where all c_j are nonnegative and the measurable sets $C_1,\dots,C_k$ form a (not necessarily strictly) decreasing chain. Such representations are called *comonotone representations* since the functions $c_1 \cdot \mathbf{1}_{C_1}, \dots, c_k \cdot \mathbf{1}_{C_k}$ are pairwise comonotone. A comonotone representation (7) has a minimal number of summands. Because of [4], for each comonotone representation of s and for each monotone measure $m \in \mathcal{M}^{X,\mathcal{A}}$ we get

$$\bigoplus_{j=1}^{k} c_j \otimes m(C_j) = \bigoplus_{i=1}^{n} b_i \otimes m(A_i).$$

This fact implies the monotonicity of $\mathbf{I}^{\text{simple}}_{\oplus,\otimes}$ on $\mathcal{F}^{(X,\mathcal{A})}_{\text{simple}}$.

If the underlying set X is finite (e.g., a set of criteria), then each $f \in \mathcal{F}^{(X,\mathcal{A})}$ is simple and, therefore, $\mathbf{I}^{\text{simple}}_{\oplus,\otimes}$ is a well-defined universal integral in this context. Note that this is the case in the majority of applications in engineering and social choice. The following proposition discusses the general case, introducing a universal integral on an arbitrary measurable space $(X,\mathcal{A})$; its proof can be found in [38].

Proposition 2. *Let $(\oplus, \otimes)$ be an arbitrary integral operation pair. The function $\mathbf{I}_{\oplus,\otimes}\colon \mathcal{D}_{[0,\infty]} \to [0,\infty]$ given by*

$$\mathbf{I}_{\oplus,\otimes}(m,f) = \sup\{\mathbf{I}^{simple}_{\oplus,\otimes}(\mu,s) \mid (\mu,s) \in \mathcal{D}_{[0,\infty]}, s \text{ simple}, h^{(\mu,s)} \le h^{(m,f)}\} \tag{9}$$

is a universal integral which is an extension of $\mathbf{I}^{simple}_{\oplus,\otimes}$ in (8).

Note that the Choquet-like integrals in [46] are special cases of universal integrals of the type $\mathbf{I}_{\oplus,\otimes}$, with $(\oplus,\otimes)$ being an appropriate integral operation pair. Here are some more observations based on Proposition 2:

(i) For each pseudo-multiplication $\otimes$ on $[0,\infty]$, the pair $(\sup,\otimes)$ is an integral operation pair and we have $\mathbf{I}_{\sup,\otimes} = \mathbf{I}_{\otimes}$ (implying $\mathbf{Su} = \mathbf{I}_{\sup,\mathrm{Min}}$).

(ii) The Choquet integral is related to the pair $(+,\mathrm{Prod})$, i.e., $\mathbf{Ch} = \mathbf{I}_{+,\mathrm{Prod}}$.

(iii) For each $p \in\,]0,\infty[$, let us define the pseudo-addition $+_p\colon\ [0,\infty]^2 \to [0,\infty]$ by

$$a +_p b = (a^p + b^p)^{1/p}.$$

The pair $(+_p,\mathrm{Prod})$ is an integral operation pair, and we have

$$\mathbf{I}_{+_p,\mathrm{Prod}}(m,f) = (\mathbf{Ch}(m^p, f^p))^{1/p},$$

i.e., $\mathbf{I}_{+_p,\mathrm{Prod}}$ can be derived from the Choquet integral $\mathbf{Ch}$ (although by a non-standard transformation applying the function $x \mapsto x^p$ to both arguments of the pair (m,f)). Obviously, $\mathbf{I}_{+_p,\mathrm{Prod}}$ is neither join- nor meet-homogeneous, and it is comonotone additive if and only if $p = 1$.

(iv) If the pseudo-multiplication satisfies the two mild continuity assumptions (CRB) and (CLZ) in [56] then Theorem 3(b) in [56] shows that the only pseudo-additions are the maximum or a strict t-conorm on $[0,\infty]$ (observe that in equation (59) on page 407 in [56] the name of the condition should be (Z) rather than (C)). This non-trivial result also shows that the left-distributivity assumption restricts the pseudo-additions and pseudo-multiplications on finite, i.e., bounded intervals, since with the usual multiplication as pseudo-multiplicaton and the restricted addition (on finite intervals) as pseudo-addition we have no left-distributivity (see [56, Example 1]).

4 Restriction to the Unit Interval

For monotone measures $m \in \mathcal{M}_1^{(X,\mathcal{A})}$ (i.e., satisfying $m(X) = 1$) and functions $f \in \mathcal{F}^{(X,\mathcal{A})}$ satisfying $\mathrm{Ran}(f) \subseteq [0,1]$ (in which case we shall write shortly $f \in \mathcal{F}_{[0,1]}^{(X,\mathcal{A})}$), the known universal integrals are related to pseudo-multiplications $\otimes$ with neutral element 1. In such a case the restriction of a universal integral to the class

$$\mathcal{D}_{[0,1]} = \bigcup_{(X,\mathcal{A})\in\mathcal{S}} \mathcal{M}_1^{(X,\mathcal{A})} \times \mathcal{F}_{[0,1]}^{(X,\mathcal{A})}$$

will be called a *universal integral on the scale* $[0,1]$ or simply a *fuzzy integral*. Observe that, in this case, only the restriction of the pseudo-multiplication $\otimes$ to

$[0,1]^2$ (again denoted by $\otimes$) is exploited, which is also called a semicopula or a conjunctor or a t-seminorm.

Recall that a *semicopula* (also called a *conjunctor* or a *t-seminorm*) [3, 12, 65] is a binary operation $\otimes\colon [0,1]^2 \to [0,1]$ which is non-decreasing in both components, has 1 as neutral element and satisfies $a \otimes b \leq \min(a,b)$ for all $(a,b) \in [0,1]^2$.

If, moreover, the pseudo-multiplication $\otimes\colon [0,1]^2 \to [0,1]$ is associative and commutative, then it is called a *triangular norm* (*t-norm* for short, see [45, 59] and also [1, 35, 60]). The four prototypical t-norms are

(i) $T_{\mathbf{M}}$, i.e., the restriction of Min to $[0,1]^2$, which is the greatest semicopula and the background of the original Sugeno integral as introduced in [66];
(ii) $T_{\mathbf{P}}$, i.e., the restriction of Prod to $[0,1]^2$, background of the original Shilkret integral as introduced in [61] and of the Choquet integral on the scale $[0,1]$;
(iii) $T_{\mathbf{L}}$ given by $T_{\mathbf{L}}(a,b) = \max(a+b-1,0)$ (*Łukasiewicz t-norm*);
(iv) $T_{\mathbf{D}}$ given by $T_{\mathbf{D}}(a,b) = 0$ if $(a,b) \in [0,1[^2$ and $T_{\mathbf{D}}(a,b) = a \cdot b$ otherwise, which is the smallest semicopula (*drastic product*).

Other important special semicopulas are the *copulas* introduced in [63] (see also [1, 51]), which are joint distribution functions restricted to $[0,1]^2$ of two-dimensional random vectors whose marginals are uniformly distributed on $[0,1]$. In the set of semicopulas, copulas C are characterized by the 2-*increasing* property, i.e., for all $0 \leq a \leq b \leq 1$ and $0 \leq c \leq d \leq 1$ we have $C(a,c) + C(b,d) - C(b,c) - C(a,d) \geq 0$. Observe that the left-hand side of this inequality is just the probability of the rectangle $[a,b] \times [c,d]$ related to the two-dimensional probability distribution described by the copula C.

Of the four prototypical t-norms mentioned above, only $T_{\mathbf{D}}$ is not a copula, the others characterize three special cases of stochastic dependence of the marginal random variables [51, 60]: $T_{\mathbf{M}}$ describes total positive dependence, $T_{\mathbf{P}}$ independence, and $T_{\mathbf{L}}$ total negative dependence.

All constructions concerning universal integrals on the $[0,\infty]$ scale based on a fixed pseudo-multiplication on $[0,\infty]$ can be applied also to universal integrals on the $[0,1]$ scale, starting from a fixed semicopula. As an example, the smallest and the greatest universal integral $\mathbf{I}_{\circledast}$ and $\mathbf{I}^{\circledast}$ on the $[0,1]$ scale related to the semicopula $\circledast$ are given by, respectively:

$$\mathbf{I}_{\circledast}(m,f) = \sup\{t \circledast m(\{f \geq t\}) \mid t \in [0,1]\},$$
$$\mathbf{I}^{\circledast}(m,f) = \operatorname{essup}_m f \circledast \sup\{m(\{f \geq t\}) \mid t \in]0,1]\}.$$

Note that, for a fixed strict t-norm T, the corresponding universal integral $\mathbf{I}_T$ is the so-called *Sugeno-Weber integral* [72]. Moreover, for a general semicopula T, the corresponding universal integral $\mathbf{I}_{\circledast}$ was called a *seminormed integral* in [65]. Note that for each semicopula $\circledast$ which is different from the minimum, the resulting universal integral $\mathbf{I}_{\circledast}$ is join-homogeneous, but neither meet-homogeneous nor comonotone additive.

Table 1. Special universal integrals

Integrals on $[0,\infty]$	**Corresponding universal integral**
Sugeno integral [66]	$\mathbf{I}_{\mathrm{Min}}$
Choquet integral [10]	$\mathbf{I}_{+,\mathrm{Prod}}$
Shilkret integral [61]	$\mathbf{I}_{\mathrm{Prod}}$
Choquet-like integral [46]	$\mathbf{I}_{\oplus,\otimes}$ [$(\oplus,\otimes)$ *appropriate integral operation pair*]
Integrals on $[0,1]$	
Sugeno-Weber integral [72]	$\mathbf{I}_T$ [T *strict t-norm*]
seminormed integral [65]	$\mathbf{I}_{\circledast}$ [$\circledast$ *semicopula*]
Imaoka integral [26, 27]	$\mathbf{K}_C$ [C *Frank copula*]
Sugeno integral [66]	$\mathbf{K}_{T_{\mathbf{M}}}$
Choquet integral [10]	$\mathbf{K}_{T_{\mathbf{P}}}$
opposite Sugeno integral [26, 27]	$\mathbf{K}_{T_{\mathbf{L}}}$

The fact that copulas are special semicopulas allows us to give another construction of universal integrals, following the computation of the standard expected value in probability theory and inspired by [26, 27] (compare also [36]) (for a proof see [38]):

Proposition 3. *Let* $C\colon [0,1]^2 \to [0,1]$ *be a copula and define* $\mathbf{K}_C\colon \mathcal{D}_{[0,1]} \to [0,1]$ *by*

$$\mathbf{K}_C(m,f) = P_C(\{(x,y) \in]0,1]^2 \mid y < m(\{f \geq x\})\}), \tag{10}$$

where P_C *is the probability measure on* $\mathcal{B}([0,1]^2)$ *induced by C, i.e., for all* $(a,b) \in [0,1]^2$ *we have* $P_C(]0,a[\times]0,b[) = C(a,b)$*. Then* $\mathbf{K}_C$ *is a universal integral on the scale* $[0,1]$.

Let us conclude this survey with some remarks about copula-based integrals on $[0,1]$ (for more properties of discrete integrals of this type see [33]):

(i) The universal integral $\mathbf{K}_{T_{\mathbf{P}}}$ on $[0,1]$ is just the Choquet integral on $[0,1]$. The copula $T_{\mathbf{P}}$ indicates that in this case the values of the underlying monotone measure and of the integrand act independently from each other. In particular, if $m(A) > 0$ and $f_t = t \cdot \mathbf{1}_A$ then $\mathbf{K}_{T_{\mathbf{P}}}(m,f_t)$ is strictly increasing in t. On the other hand, $\mathbf{K}_{T_{\mathbf{M}}}$ is exactly the Sugeno integral as introduced in [66]. The copulas $T_{\mathbf{M}}$ and $T_{\mathbf{L}}$ indicate maximal positive and maximal negative dependence between the values of the underlying monotone measure and of the integrand, respectively. As a consequence, for m, A and f_t as above, $\mathbf{K}_{T_{\mathbf{M}}}(m,f_t)$ is strictly increasing in t only if $t < m(A)$, while $\mathbf{K}_{T_{\mathbf{L}}}(m,f_t)$ is strictly increasing in t only if $t > 1 - m(A)$.

(ii) If X is a finite set, i.e., if $X = \{1,2,\ldots,n\}$ with $n \in \mathbb{N}$ and $\mathcal{A} = 2^X$, each function $f \in \mathcal{F}_{[0,1]}^{(X,\mathcal{A})}$ corresponds to the n-tuple $(f(1), f(2), \ldots, f(n)) \in [0,1]^n$. Let σ be a permutation of X with $f(\sigma(1)) \leq f(\sigma(2)) \leq \cdots \leq f(\sigma(n))$, and put $A_i = \{\sigma(i), \ldots, \sigma(n)\}$. Then, using the conventions $A_{n+1} = \emptyset$ and $f(\sigma(0)) = 0$, we obtain two equivalent expressions for $\mathbf{K}_C(m,f)$, namely,

$$\mathbf{K}_C(m,f) = \sum_{i=1}^{n}(C(f(\sigma(i)),m(A_i)) - C(f(\sigma(i)),m(A_{i+1})))$$
$$= \sum_{i=1}^{n}(C(f(\sigma(i)),m(A_i)) - C(f(\sigma(i-1)),m(A_i))).$$

It is possible to introduce universal integrals on an arbitrary fixed scale $[0,e]$ with $e \in]0,\infty]$, considering, for a fixed measurable space $(X,\mathcal{A})$, monotone measures in $\mathcal{M}_e^{(X,\mathcal{A})}$ and functions $f \in \mathcal{F}^{(X,\mathcal{A})}$ with $\mathrm{Ran}(f) \subseteq [0,e]$, and working with a pseudo-multiplication $\otimes$ with neutral element e (in fact, only the restriction of $\otimes$ to $[0,e]^2$ is needed). However, for each universal integral $\mathbf{U}^{[0,e]}$ on $[0,e]$ there exists a universal integral $\mathbf{I}$ on the scale $[0,1]$ such that:

$$\mathbf{U}^{[0,e]}(m,f) = \begin{cases} e \cdot \mathbf{I}\left(\frac{m}{e},\frac{f}{e}\right) & \text{if } e < \infty, \\ \frac{\mathbf{I}\left(\frac{m}{m+1},\frac{f}{f+1}\right)}{1-\mathbf{I}\left(\frac{m}{m+1},\frac{f}{f+1}\right)} & \text{if } e = \infty. \end{cases} \tag{11}$$

Conversely, if $\mathbf{I}$ is a universal integral on the scale $[0,1]$ then $\mathbf{U}^{[0,e]}$ constructed via (11) is a universal integral on $[0,e]$.

If we start with $\mathbf{I} = \mathbf{K}_{T_\mathbf{M}}$ (i.e., with the Sugeno integral on the scale $[0,1]$), then formula (11) always leads to the Sugeno integral $\mathbf{Su}^{[0,e]}$, due to the fact that the restriction of Min to $[0,e]^2$ has neutral element e. On the other hand, if $\mathbf{I} = \mathbf{K}_{T_\mathbf{P}}$ (i.e., the Choquet integral on the scale $[0,1]$), then formula (11) leads to a universal integral which is different from the Choquet integral on $[0,e]$, whenever $e \neq 1$.

References

1. Alsina, C., Frank, M.J., Schweizer, B.: Associative functions: Triangular norms and copulas. World Scientific, Singapore (2006)
2. Aumann, R.J., Shapley, L.S.: Values of Non-Atomic Games. Princeton University Press, Princeton (1974)
3. Bassan, B., Spizzichino, F.: Relations among univariate aging, bivariate aging and dependence for exchangeable lifetimes. J. Multivariate Anal. 93, 313–339 (2005)
4. Benvenuti, P., Mesiar, R., Vivona, D.: Monotone set functions-based integrals. In: Pap [54], pp. 1329–1379 (2002)
5. Butnariu, D.: Additive fuzzy measures and integrals. I. J. Math. Anal. Appl. 93, 436–452 (1983)
6. Butnariu, D., Klement, E.P.: Triangular norm-based measures and their Markov kernel representation. J. Math. Anal. Appl. 162, 111–143 (1991)
7. Butnariu, D., Klement, E.P.: Triangular Norm-Based Measures and Games with Fuzzy Coalitions. Kluwer Academic Publishers, Dordrecht (1993)
8. Calvo, T., Kolesárová, A., Komorníková, M., Mesiar, R.: Aggregation operators: properties, classes and construction methods. In: Calvo et al [9], pp. 3–104 (2002)
9. Calvo, T., Mayor, G., Mesiar, R. (eds.): Aggregation operators. New trends and applications. Physica-Verlag, Heidelberg (2002)

10. Choquet, G.: Theory of capacities. Ann. Inst. Fourier (Grenoble) 5, 131–292 (1953–1954)
11. Denneberg, D.: Non-additive measure and integral. Kluwer Academic Publishers, Dordrecht (1994)
12. Durante, F., Sempi, C.: Semicopulæ. Kybernetika (Prague) 41, 315–328 (2005)
13. Figueira, J., Greco, S., Ehrgott, M. (eds.): Multiple criteria decision analysis. State of the art surveys. Springer, New York (2005)
14. Frank, M.J.: On the simultaneous associativity of $F(x,y)$ and $x+y-F(x,y)$. Aequationes Math. 19, 194–226 (1979)
15. Grabisch, M.: Fuzzy integral in multicriteria decision making. Fuzzy Sets and Systems 69, 279–298 (1995)
16. Grabisch, M., Labreuche, C.: The symmetric and asymmetric Choquet integrals on finite spaces for decision making. Statist. Papers 43, 37–52 (2002)
17. Grabisch, M., Labreuche, C.: A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid. 4OR 6, 1–44 (2008)
18. Grabisch, M., Marichal, J.-L., Mesiar, R., Pap, E.: Aggregation Functions. Cambridge University Press, Cambridge (2009)
19. Grabisch, M., Murofushi, T., Sugeno, M.: Fuzzy measures and integrals. Theory and applications. Physica-Verlag, Heidelberg (2000)
20. Grabisch, M., Nicolas, J.-M.: Classification by fuzzy integral: performance and tests. Fuzzy Sets and Systems 65, 255–271 (1994)
21. Grabisch, M., Raufaste, E.: An empirical study of statistical properties of Choquet and Sugeno integrals. IEEE Trans. Fuzzy Systems (in press)
22. Grabisch, M., Roubens, M.: Application of the Choquet integral in multicriteria decision making. In: Grabisch et al [19], pp. 348–374 (2000)
23. Greco, G.H.: Monotone integrals. Rend. Sem. Mat. Univ. Padova 57, 149–166 (1977)
24. Halmos, P.R.: Measure theory. Van Nostrand Reinhold, New York (1950)
25. Ichihashi, H., Tanaka, M., Asai, K.: Fuzzy integrals based on pseudo-additions and multiplications. J. Math. Anal. Appl. 50, 74–79 (1988)
26. Imaoka, H.: On a subjective evaluation model by a generalized fuzzy integral. Internat. J. Uncertain. Fuzziness Knowledge-Based Systems 5, 517–529 (1997)
27. Imaoka, H.: Comparison between three fuzzy integrals. In: Grabisch et al [19], pp. 273–286 (2000)
28. Keller, J.M., Gader, P., Tahani, H., Chiang, J.-H., Mohamed, M.: Advances in fuzzy integration for pattern recognition. Fuzzy Sets and Systems 65, 273–283 (1994)
29. Klement, E.P.: Fuzzy σ-algebras and fuzzy measurable functions. Fuzzy Sets and Systems 4, 83–93 (1980)
30. Klement, E.P.: Characterization of finite fuzzy measures using Markoff-kernels. J. Math. Anal. Appl. 75, 330–339 (1980)
31. Klement, E.P.: Construction of fuzzy σ-algebras using triangular norms. J. Math. Anal. Appl. 85, 543–565 (1982)
32. Klement, E.P.: Characterization of fuzzy measures constructed by means of triangular norms. J. Math. Anal. Appl. 86, 345–358 (1982)
33. Klement, E.P., Kolesárová, A., Mesiar, R., Stupňanová, A.: Lipschitz continuity of discrete universal integrals based on copulas. J. Uncertain. Fuzziness Knowledge-Based Systems 18, 39–52 (2010)
34. Klement, E.P., Lowen, R., Schwyhla, W.: Fuzzy probability measures. Fuzzy Sets and Systems 5, 21–30 (1981)
35. Klement, E.P., Mesiar, R., Pap, E.: Triangular norms. Kluwer Academic Publishers, Dordrecht (2000)

36. Klement, E.P., Mesiar, R., Pap, E.: Measure-based aggregation operators. Fuzzy Sets and Systems 142, 3–14 (2004)
37. Klement, E.P., Mesiar, R., Pap, E.: A universal integral. In: Proceedings EUSFLAT 2007, Ostrava, vol. I, pp. 253–256 (2007)
38. Klement, E.P., Mesiar, R., Pap, E.: A universal integral as common frame for Choquet and Sugeno integral. IEEE Trans. Fuzzy Systems 18, 178–187 (2010)
39. Klement, E.P., Ralescu, D.A.: Nonlinearity of the fuzzy integral. Fuzzy Sets and Systems 11, 309–315 (1983)
40. König, H.: Measure and Integration. An Advanced Course in Basic Procedures and Applications. Springer, Berlin (1997)
41. Kroupa, T.: Application of the Choquet integral to measures of information in possibility theory. International Journal of Intelligent Systems 21, 349–359 (2006)
42. Marichal, J.-L.: On Sugeno integral as an aggregation function. Fuzzy Sets and Systems 114, 347–365 (2000)
43. Marichal, J.-L.: An axiomatic approach of the discrete Sugeno integral as a tool to aggregate interacting criteria in a qualitative framework. IEEE Trans. Fuzzy Systems 9, 164–172 (2001)
44. Mendez-Vazquez, A., Gader, P., Keller, J.M., Chamberlin, K.: Minimum classification error training for Choquet integrals with applications to landmine detection. IEEE Trans. Fuzzy Systems 16, 225–238 (2008)
45. Menger, K.: Statistical metrics. Proc. Nat. Acad. Sci. U.S.A. 8, 535–537 (1942)
46. Mesiar, R.: Choquet-like integrals. J. Math. Anal. Appl. 194, 477–488 (1995)
47. Mesiar, R., Mesiarová, A.: Fuzzy integrals—What are they? International Journal of Intelligent Systems 23, 199–212 (2008)
48. Mesiar, R., Mesiarová, A.: Fuzzy integrals and linearity. Internat. J. Approx. Reason. 47, 352–358 (2008)
49. Murofushi, T., Sugeno, M.: Fuzzy t-conorm integrals with respect to fuzzy measures: generalization of Sugeno integral and Choquet integral. Fuzzy Sets and Systems 42, 57–71 (1991)
50. Negoita, C.V., Ralescu, D.A.: Representation theorems for fuzzy concepts. Kybernetes 4, 169–174 (1975)
51. Nelsen, R.B.: An introduction to copulas, 2nd edn. Lecture Notes in Statistics , vol. 139. Springer, New York (2006)
52. Pap, E.: An integral generated by a decomposable measure. Univ. u Novom Sadu Zb. Rad. Prirod.-Mat. Fak. Ser. Mat. 20, 135–144 (1990)
53. Pap, E.: Null-additive set functions. Kluwer Academic Publishers, Dordrecht (1995)
54. Pap, E. (ed.): Handbook of measure theory. Elsevier Science, Amsterdam (2002)
55. Ralescu, D., Adams, G.: The fuzzy integral. J. Math. Anal. Appl. 75, 562–570 (1980)
56. Sander, W., Siedekum, J.: Multiplication, distributivity and fuzzy-integral. I. Kybernetika (Prague) 41, 397–422 (2005)
57. Schmeidler, D.: Integral representation without additivity. Proc. Amer. Math. Soc. 97, 255–261 (1986)
58. Schmeidler, D.: Subjective probability and expected utility without additivity. Econometrica 57, 571–587 (1989)
59. Schweizer, B., Sklar, A.: Statistical metric spaces. Pacific J. Math. 10, 313–334 (1960)
60. Schweizer, B., Sklar, A.: Probabilistic metric spaces. North-Holland, New York (1983); reprinted: Dover, Mineola (2005)
61. Shilkret, N.: Maxitive measure and integration. Indag. Math. 33, 109–116 (1971)
62. Šipoš, J.: Integral with respect to a pre-measure. Math. Slovaca 29, 141–155 (1979)

63. Sklar, A.: Fonctions de répartition à n dimensions et leurs marges. Publ. Inst. Statist. Univ. Paris 8, 229–231 (1959)
64. Struk, P.: Extremal fuzzy integrals. Soft Computing 10, 502–505 (2006)
65. Suárez García, F., Gil Álvarez, P.: Two families of fuzzy integrals. Fuzzy Sets and Systems 18, 67–81 (1986)
66. Sugeno, M.: Theory of fuzzy integrals and its applications. Ph.D. thesis, Tokyo Institute of Technology (1974)
67. Sugeno, M., Murofushi, T.: Pseudo-additive measures and integrals. J. Math. Anal. Appl. 122, 197–222 (1987)
68. Torra, V., Narukawa, Y.: The h-index and the number of citations: Two fuzzy integrals. IEEE Trans. Fuzzy Systems 16, 795–797 (2008)
69. Vitali, G.: Sulla definizione di integrale delle funzioni di una variabile. Ann. Mat. Pura Appl. 2, 111–121 (1925)
70. Wang, Z., Klir, G.J.: Fuzzy measure theory. Plenum Press, New York (1992)
71. Wang, Z., Klir, G.J.: Generalized measure theory. Springer, New York (2009)
72. Weber, S.: Two integrals and some modified versions: critical remarks. Fuzzy Sets and Systems 20, 97–105 (1986)
73. Yoneda, M., Fukami, S., Grabisch, M.: Interactive determination of a utility function represented as a fuzzy integral. Inform. Sci. 71, 43–64 (1993)
74. Zadeh, L.A.: Fuzzy sets. Inform. and Control 8, 338–353 (1965)
75. Zadeh, L.A.: Probability measures of fuzzy events. J. Math. Anal. Appl. 23, 421–427 (1968)

On Virtues of Many-Valued (Fuzzy) Type Theories

Vilém Novák

Institute for Research and Applications of Fuzzy Modeling,
University of Ostrava 30. dubna 22,
701 03 Ostrava 1, Czech Republic
e-mail: Vilem.Novak@osu.cz

Abstract. In this paper, we deal with the fuzzy type theory (FTT) — a higher-order fuzzy logic. There are several kinds of this logic depending on the chosen structure of truth values. Higher-order (fuzzy) logic is still not fully appreciated despite its high explicative power. Our goal is to point out several great virtues of it to convince the reader that this logic is worth of studying and has a potential for many applications. After brief presentation of the main algebraic structures of truth values convenient for FTT, we discuss several virtues of the latter, namely: (1) FTT has a simple and highly uniform syntax, (2) Semantics of FTT is based on a small collection of well-established ideas, (3) FTT is a highly expressive logic and (4) There are practical extensions of FTT that can be effectively implemented.

Keywords: Residuated lattice, EQ-algebra, mathematical fuzzy logic, fuzzy type theory, evaluative linguistic expressions, intermediate quantifiers.

1 Introduction

Mathematical fuzzy logic is a well established formal tool which can be applied in modeling of human reasoning affected by the vagueness phenomenon. The latter is captured via degree theoretical approach. Besides various kinds of propositional and first-order calculi, also higher-order fuzzy logic calculi have been developed. In analogy with classical logic they are called *fuzzy type theories* (FTT).

Fuzzy type theory, which is a generalization of classical type theory (cf. [1]), was introduced by V. Novák in [18]. The generalization consists especially in replacement of the axiom stating "there are just two truth values" by a sequence of axioms characterizing the structure of the algebra of truth values. This is a lattice with several additional properties. The fundamental class of algebras of truth values is formed by *MTL-algebras* which are prelinear *residuated lattices*. Another very general class of algebras especially convenient as the algebras of truth values for FTT is formed by EQ-algebras

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 53–69.
springerlink.com

which are lower semilattices with top element, the operation of fuzzy equality and a binary operation of fusion.

The syntax of FTT is generalization of the classical lambda-calculus, which differs from the classical one by definition of additional special connectives, and by logical axioms. The fundamental connective in FTT is that of a *fuzzy equality* $\equiv$, which is interpreted by a reflexive, symmetric and $\otimes$-transitive binary fuzzy relation ($\otimes$ is the fusion operation). The generalized completeness theorem has been proved for all kinds of FTT (for the details see [17, 18, 24]).

Why should type theory be made fuzzy? The FTT provides model of some deep manifestations of the vagueness phenomenon (including higher order vagueness). Furthermore, semantics of concepts and natural language expressions is formalized using TT. Hence, replacing it by FTT could enable to include vagueness in the developed models of natural language semantics and bring the formal theory of commonsense reasoning closer to the human way of thinking. Other interesting application is to establish foundations of the whole "fuzzy" mathematics which has been initiated as a special program by P. Cintula and L. Běhounek in [2]. The expressive power of FTT makes all these tasks easier.

In [6], W. Farmer discussed seven virtues of classical simple type theory (STT), namely:

(i) STT has a simple and highly uniform syntax.
(ii) The semantics of STT is based on a small collection of well-established ideas.
(iii) STT is a highly expressive logic.
(iv) STT admits categorical theories of infinite structures.
(v) There is a simple, elegant, and powerful proof system for STT.
(vi) Techniques of first-order model theory can be applied to STT; distinction between standard and nonstandard models is illuminated.
(vii) There are practical extensions of STT that can be effectively implemented.

In this paper, we will provide a brief overview of fuzzy type theories and demonstrate that they share the virtues of STT. Namely, we will show that virtues (i)–(iii) and (vii) are also those of FTT while virtues (iv)–(vi) are quite specific and might have different meaning for FTT.

2 Truth Values for FTT

The basic structure of truth values in fuzzy logics is that of residuated lattice that is integral, commutative, bounded, residuated lattice

$$\mathscr{L} = \langle L, \vee, \wedge, \otimes, \rightarrow, \mathbf{0}, \mathbf{1} \rangle \tag{1}$$

such that $\mathscr{L} = \langle L, \vee, \wedge, \mathbf{0}, \mathbf{1}\rangle$ is a lattice with $\mathbf{0}, \mathbf{1}$, $\langle L, \otimes, \mathbf{1}\rangle$ is a commutative monoid and the following adjunction property holds:

$$a \otimes b \leq c \quad \text{iff} \quad a \leq b \rightarrow c.$$

(note that this is algebraic formulation of the rule of modus ponens).

Further assumed properties of algebras of truth degrees are the following:

(i) prelinearity $(a \rightarrow b) \vee (b \rightarrow a) = \mathbf{1}$,
(ii) divisibility $a \otimes (a \rightarrow b) = a \wedge b$,
(iii) double negation $\neg\neg a = a$.

We distinguish the following classes of special residuated lattices:

(i) *MTL-algebra* is a residuated lattice with prelinearity.
(ii) *IMTL-algebra* is a residuated lattice with prelinearity and double negation.
(iii) *BL-algebra* is a residuated lattice with prelinearity and divisibility.
(iv) *MV-algebra* is a residuated lattice with prelinearity, divisibility and double negation.

For FTT, it is important to consider the also the *delta operation* which, in case that $\mathscr{L}$ is linearly ordered, has the following definition[1]:

$$\Delta(a) = \begin{cases} \mathbf{1} & \text{if } a = \mathbf{1}, \\ \mathbf{0} & \text{otherwise.} \end{cases} \tag{2}$$

A very general class of algebras especially convenient as the algebras of truth values for FTT is formed by *EQ-algebras*. These are algebras

$$\mathscr{E} = \langle E, \wedge, \otimes, \sim, \mathbf{1}\rangle$$

of type (2, 2, 2, 0), where

(E1) $\langle E, \wedge\rangle$ is a $\wedge$-semilattice with the top element $\mathbf{1}$,
(E2) $\langle L, \otimes, \mathbf{1}\rangle$ is a commutative monoid and $\otimes$ is isotone w.r.t. $\leq$ ($a \leq b$ iff $a \wedge b = a$),
(E3) $a \sim a = \mathbf{1}$, (reflexivity)
(E4) $((a \wedge b) \sim c) \otimes (d \sim a) \leq c \sim (d \wedge b)$, (substitution)[2].
(E5) $(a \sim b) \otimes (c \sim d) \leq (a \sim c) \sim (b \sim d)$, (congruence)
(E6) $(a \wedge b \wedge c) \sim a \leq (a \wedge b) \sim a$, (monotonicity)
(E7) $a \otimes b \leq a \sim b$. (boundedness)

[1] The Δ-operation is, in general, determined by 6 special properties. For the details, see [9, 18, 24].

[2] This property corresponds to the Leibniz rule of indiscernibility of identicals.

The following are special definitions in EQ-algebras:

(i) $\tilde{a} = a \sim \mathbf{1}$.

(ii) $a \rightarrow b = (a \wedge b) \sim a$. (implication)

(iii) The multiplication $\otimes$ is $\rightarrow$-isotone if

$$a \rightarrow b = \mathbf{1} \quad \text{implies} \quad a \otimes c \rightarrow b \otimes c = \mathbf{1}.$$

(iv) If $\mathscr{E}$ contains $\mathbf{0}$ then $\neg a = a \sim \mathbf{0}$. (negation)

(v) $a \leftrightarrow b = (a \rightarrow b) \wedge (b \rightarrow a)$. (biimplication)

(vi) $a \overset{\circ}{\leftrightarrow} b = (a \rightarrow b) \otimes (b \rightarrow a)$. (weak biimplication)

The basic properties of EQ-algebras are summarized in the following theorem:

Theorem 1. *(a)* $a \sim b = b \sim a$, *(symmetry)*
(b) $(a \sim b) \otimes (b \sim c) \leq (a \sim c)$, *(transitivity)*
(c) $(a \rightarrow b) \otimes (b \rightarrow c) \leq a \rightarrow c$, *(transitivity of implication)*
(d) If $a \leq b \rightarrow c$, *then* $a \otimes b \leq \tilde{c}$. *(semi-adjunction)*

An EQ-algebra is:

(i) *separated* if $a \sim b = \mathbf{1}$ iff $a = b$.

(ii) *good* if $a \sim \mathbf{1} = a$,

(iii) *spanned* if $\tilde{\mathbf{0}} = \mathbf{0}$,

(iv) *involutive* if $\neg\neg a = a$, (IEQ-algebra)

(v) *residuated* if $(a \otimes b) \wedge c = a \otimes b$ iff $a \wedge ((b \wedge c) \sim b) = a$,

(vi) *complete* if it is $\wedge$-semilattice complete,

(vii) *ℓEQ-algebra* if it is lattice ordered and the following is fulfiled:

(E10) $((a \vee b) \sim c) \otimes (d \sim a) \leq ((d \vee b) \sim c)$.

All structures of truth values considered in FTT must contain also the delta operation (2). It should be noted that each residuated lattice gives rise to an EQ-algebra but there are EQ-algebras which are not residuated.

3 Virtues of Fuzzy Type Theory

3.1 Virtue 1: FTT has a Simple and Highly Uniform Syntax

The syntax of FTT is a slight generalization of the syntax of classical type theory — the λ-calculus — where the latter has been introduced by A. Church in [3]. Therefore, this virtue is in FTT fully appreciated too.

Types and formulas

A type is a symbol using which various kinds of formulas are characterized. The set of types *Types* is constructed recursively from the elementary types o (truth values) and ϵ (objects):

(i) $\epsilon, o \in \mathit{Types}$,
(ii) If $\alpha, \beta \in \mathit{Types}$ then $(\alpha\beta) \in \mathit{Types}$[3].

The language J of FTT consists of variables and constants of specific types. Formulas are defined as follows:

(i) If $x_\alpha \in J$ is a variable then x_α is a formula of type α.
(ii) If $c_\alpha \in J$ is a constant then c_α is a formula of type α.
(iii) If $B_{\beta\alpha}$ and A_α are formulas then $(B_{\beta\alpha} A_\alpha)$ is a formula of type β.
(iv) If A_β is a formula and $x_\alpha \in J$ a variable then $\lambda x_\alpha\, A_\beta$ is a formula of type $\beta\alpha$[4].

Formulas A_o are propositions. To reduce the burden of subscripts we sometimes write $A \in \mathit{Form}_\alpha$ to express that A is a formula of type α and then omit the subscript α at A.

Fuzzy equality

This is the crucial concept in FTT and has important role in the interpretation (see below). In syntax, it is determined by a special formula (constant)

$$\mathbf{E}_{(o\alpha)\alpha}.$$

Then, the fuzzy equality is defined by

$$\equiv\ :=\ \lambda x_\alpha \lambda y_\alpha (\mathbf{E}_{(o\alpha)\alpha}\, y_\alpha) x_\alpha.$$

We write $(x_o \equiv y_o)$ instead of $(\equiv y_o)x_o$. Similarly we write $(A_\alpha \equiv B_\alpha)$. Note that these are formulas of type o.

Logical axioms of IMTL-FTT

The structure of truth values is an IMTL$_\Delta$-algebra. Furthermore, we defined several special formulas, for example representation of *truth* and *falsity:*

$$\top := (\lambda x_o\, x_o \equiv \lambda x_o\, x_o) \qquad \bot := (\lambda x_o\, x_o \equiv \lambda x_o\, \top).$$

Moreover, all connectives in FTT are formulas, e.g.

$$\neg := \lambda x_o (\bot \equiv x_o), \qquad \text{(negation)}$$
$$\Rightarrow\ := \lambda x_o (\lambda y_o ((x_o \wedge y_o) \equiv x_o)). \qquad \text{(implication)}$$

There are more axioms in FTT than in classical TT because it is necessary to characterize the structure of truth values. Below, we list axioms of IMTL-FTT where the structure of truth values forms an IMTL$_\Delta$-algebra.

[3] The types are alternatively also written as $\alpha \to \beta$ instead of $\beta\alpha$. This makes, however, formulas too long and poorly arranged.

[4] Alternatively, it is possible to write $A{:}\alpha$ instead of A_α. In various papers, formulas are also called *lambda terms*.

Fundamental axioms

(FT1) $\boldsymbol{\Delta}(x_\alpha \equiv y_\alpha) \Rightarrow (f_{\beta\alpha}\, x_\alpha \equiv f_{\beta\alpha}\, y_\alpha)$
(FT2$_1$) $(\forall x_\alpha)(f_{\beta\alpha}\, x_\alpha \equiv g_{\beta\alpha}\, x_\alpha) \Rightarrow (f_{\beta\alpha} \equiv g_{\beta\alpha})$
(FT2$_2$) $(f_{\beta\alpha} \equiv g_{\beta\alpha}) \Rightarrow (f_{\beta\alpha}\, x_\alpha \equiv g_{\beta\alpha}\, x_\alpha)$
(FT3) $(\lambda x_\alpha B_\beta) A_\alpha \equiv C_\beta$

where C_β is obtained from B_β by replacing all free occurrences of x_α in it by A_α, provided that A_α is substitutable to B_β for x_α (*lambda conversion*).

(FT4) $(x_\epsilon \equiv y_\epsilon) \Rightarrow ((y_\epsilon \equiv z_\epsilon) \Rightarrow (x_\epsilon \equiv z_\epsilon))$

Equivalence axioms

(FT6) $(x_o \equiv y_o) \equiv ((x_o \Rightarrow y_o) \wedge (y_o \Rightarrow x_o))$
(FT7) $(A_o \equiv \top) \equiv A_o$

Implication axioms

(FT8) $(A_o \Rightarrow B_o) \Rightarrow ((B_o \Rightarrow C_o) \Rightarrow (A_o \Rightarrow C_o))$
(FT9) $(A_o \Rightarrow (B_o \Rightarrow C_o)) \equiv (B_o \Rightarrow (A_o \Rightarrow C_o))$
(FT10) $((A_o \Rightarrow B_o) \Rightarrow C_o) \Rightarrow (((B_o \Rightarrow A_o) \Rightarrow C_o) \Rightarrow C_o)$
(FT11) $(\neg B_o \Rightarrow \neg A_o) \equiv (A_o \Rightarrow B_o)$

Conjunction axioms

(FT12) $A_o \wedge B_o \equiv B_o \wedge A_o$
(FT13) $A_o \wedge B_o \Rightarrow A_o$
(FT14) $(A_o \wedge B_o) \wedge C_o \equiv A_o \wedge (B_o \wedge C_o)$

Delta axioms

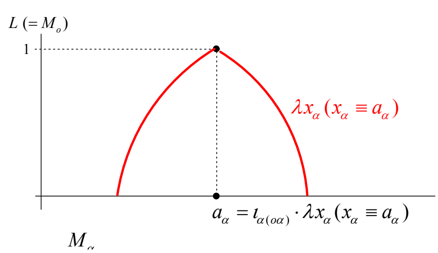
(FT5) $(g_{oo}(\boldsymbol{\Delta} x_o) \wedge g_{oo}(\neg \boldsymbol{\Delta} x_o)) \equiv (\forall y_o) g_{oo}(\boldsymbol{\Delta} y_o)$
(FT15) $\boldsymbol{\Delta}(A_o \wedge B_o) \equiv \boldsymbol{\Delta} A_o \wedge \boldsymbol{\Delta} B_o$
(FT16) $\boldsymbol{\Delta}(A_o \vee B_o) \Rightarrow \boldsymbol{\Delta} A_o \vee \boldsymbol{\Delta} B_o$

Predicate axioms

(FT17) $(\forall x_\alpha)(A_o \Rightarrow B_o) \Rightarrow (A_o \Rightarrow (\forall x_\alpha) B_o)$ x_α is not free in A_o

Axiom of descriptions

(FT18) $\iota_{\epsilon(o\epsilon)}(\mathbf{E}_{(o\epsilon)\epsilon}\, y_\epsilon) \equiv y_\epsilon$



This axiom enables to reach elements of fuzzy sets. It corresponds to the defuzzification operation in fuzzy set theory. Its meaning is schematically depicted in the picture above.

Inference rules and provability

(Rule R) *Let* $A_\alpha \equiv A'_\alpha$ *and* $B \in Form_o$*. Then infer* B' *where* B' *comes form* B *by replacing one occurrence of* A_α*, which is not preceded by* λ*, by* A'_α*.*

(Rule (N)) *Let* $A_o \in Form_o$ *be a formula. Then from* A_o *infer* $\boldsymbol{\Delta} A_o$*.*

Note that the rule (R) is the same as in classical TT. On the other hand the rules of modus ponens and generalization are derived in FTT. A theory T of FTT is a set of formulas of type o. The provability $T \vdash A_o$ is defined as usual.

Theorem 2 (Deduction theorem). *Let* T *be a theory,* $A_o \in Form_o$ *a formula. Then*

$$T \cup \{A_o\} \vdash B_o \quad \textit{iff} \quad T \vdash \boldsymbol{\Delta} A_o \Rightarrow B_o$$

holds for every formula $B_o \in Form_o$*.*

3.2 Virtue 2: Semantics of FTT Is Based on a Small Collection of Well-Established Ideas

In FTT, the algebra of truth values should be one of the following:

1. A complete linearly ordered IMTL$_\Delta$-algebra,
2. linearly ordered Łukasiewicz$_\Delta$-algebra,
3. linearly ordered BL$_\Delta$-algebra,
4. linearly ordered EQ$_\Delta$-algebra or IEQ$_\Delta$-algebra.

General frame

The semantics of FTT is defined with respect to the general frame

$$\mathscr{M} = \langle \{M_\alpha, \stackrel{\circ}{=}_\alpha \mid \alpha \in Types\}, \mathscr{E}_\Delta \rangle$$

where $\mathscr{E}_\Delta$ is the algebra of truth values (one of the above ones) and each type α is assigned a set M_α together with a *fuzzy equality*

$$\stackrel{\circ}{=}_\alpha : M_\alpha \times M_\alpha \longrightarrow L.$$

The sets are constructed as follows:

(i) M_o is the set of truth values (support of the algebra $\mathscr{E}_\Delta$),
(ii) M_ϵ is some (non-empty) set,
(iii) $M_{\beta\alpha} \subseteq M_\beta^{M_\alpha}$,
(iv) $M_{oo} \cup M_{(oo)o}$ is closed w.r.t. operations on truth values,
(v) $\stackrel{\circ}{=}_\alpha$ is a fuzzy equality on M_α:
  (a) $\stackrel{\circ}{=}_o$ is $\sim$ (or $\leftrightarrow$),
  (b) $\stackrel{\circ}{=}_\epsilon$ is given explicitly,

$$\text{(c)}\ \ [h \mathrel{\mathring{\doteq}}_{\beta\alpha} h'] = \bigwedge_{m \in M_\alpha} [h(m) \mathrel{\mathring{\doteq}}_\beta h'(m)], \qquad h, h' : M_\alpha \longrightarrow M_\beta.$$

The general frame can be schematically depicted as follows:

$$(M_o = \{a \mid a \in L\}, \leftrightarrow) \qquad (M_\epsilon = \{u \mid \varphi(u)\}, =_\epsilon)$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$(M_{oo} \subseteq \{g_{oo} \mid g_{oo} : M_o \longrightarrow M_o\}, =_{oo})$$

$$(M_{o\epsilon} \subseteq \{f_{o\epsilon} \mid f_{o\epsilon} : M_\epsilon \longrightarrow M_o\}, =_{o\epsilon})$$

$$(M_{\epsilon\epsilon} \subseteq \{f_{\epsilon\epsilon} \mid f_{\epsilon\epsilon} : M_\epsilon \longrightarrow M_\epsilon\}, =_{\epsilon\epsilon}), \ldots$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\vdots$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$(M_{\beta\alpha} \subseteq \{f_{\beta\alpha} \mid f_{\beta\alpha} : M_\alpha \longrightarrow M_\beta\}, =_{\beta\alpha})$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\vdots$$

Interpretation of formulas

Each formula of type α is in the general frame assigned and element from the set of the same type, i.e.,

$$\mathscr{M}(A_{\beta\alpha}) \in M_{\beta\alpha}.$$

It is important to assure that the interpretation of formulas must preserve the fuzzy equality.

Example 1.

- $\mathscr{M}(A_o) \in L$ is a truth value,
- $\mathscr{M}(A_{o\epsilon})$ is a fuzzy set in M_ϵ,
- $\mathscr{M}(A_{(o\epsilon)\epsilon})$ is a fuzzy relation on M_ϵ,
- $\mathscr{M}(A_{(oo)\epsilon})$ is a fuzzy set of type 2,
- $\mathscr{M}(A_{\epsilon\epsilon})$ is a function on objects.

A *model* of T is a frame $\mathscr{M}$ in which all special axioms of T are true in the degree $\mathbf{1}$. A formula A_o is true in T, $T \models A_o$, if it is true in the degree $\mathbf{1}$ in all models of T.

Theorem 3 (Completeness).

(a) A theory T of IMTL-FTT is consistent iff it has a general model $\mathscr{M}$.
(b) For every theory T of IMTL-FTT and a formula A_o

$$T \vdash A_o \quad \textit{iff} \quad T \models A_o.$$

A scheme of the main kinds of fuzzy type theories is depicted in Figure 1.

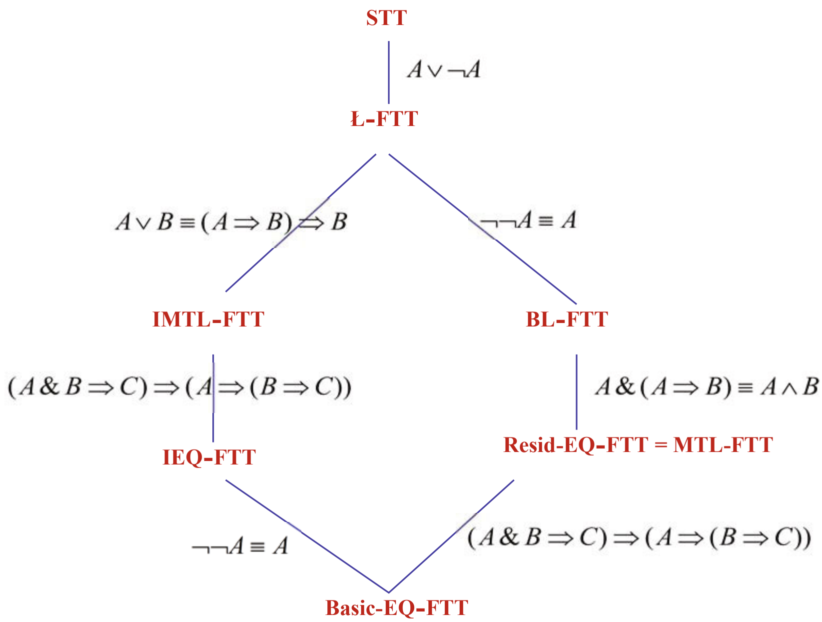


Fig. 1. A scheme of the main kinds of fuzzy type theory

3.3 Virtue 3: FTT Is a Highly Expressive Logic

We argue that FTT is a powerful logic using which it is possible to express important manifestations of the vagueness phenomenon. Recall from [20] that *vagueness* raises when trying to group together objects carrying a certain property φ. We form an *actualized* grouping of objects

$$X = \{o \mid \varphi(o)\}.$$

Then we can distinguish typical objects having φ as well as those not having it, and *borderline* objects for which it is unclear whether they have the property φ. Moreover, we also encounter imperceptible gradual change of the property φ from its presence to its non-presence. This is the continuity which is typical feature of vagueness. We argue that the most distinguished mathematical model of vagueness is provided by *fuzzy logic*. The mathematization is based on introduction of a *numerical measure of the truth that a particular object has a property in concern*. We claim that *all essential properties of vague predicates are formally expressible in FTT and so, they have a many-valued model.*

One of very popular models of vagueness is the *supervaluation theory* (cf. [12]). According to it, a sentence is precise if it is precise under any possible precisification. The theory thus introduces the following concepts: a proposition A is *supertrue* if it is true under any truth valuation and we can write

$D(A)$. It is *superfalse* if it is false under any truth valuation. Otherwise it is *undefined* and we can write $I(A)$.

These concepts can be easily expressed inside FTT. Namely, the $\boldsymbol{\Delta}$ connective corresponds to D-operator and behaves accordingly. For example $A \vdash C$ implies $\vdash \boldsymbol{\Delta} A \Rightarrow C$ as well as $\neg C \vdash \neg\boldsymbol{\Delta} A$ is provable in FTT. We can further introduce the following special unary connectives:

$$\Upsilon_{oo} \equiv \lambda z_o \cdot \neg\boldsymbol{\Delta}(\neg z_o),$$

Then the following holds in any model $\mathscr{M}$:

$$\mathscr{M}(\Upsilon z_o) = \mathbf{1} \quad \text{iff} \quad z_0 > \mathbf{0}.$$

Similarly, we can also introduce the connective characterizing general truth value:

$$\hat{\Upsilon}_{oo} \equiv \lambda z_o \cdot \neg\boldsymbol{\Delta}(z_o \vee \neg z_o) \equiv \lambda z_o \cdot \neg\boldsymbol{\Delta} z_o \wedge \neg\boldsymbol{\Delta}\neg z_o).$$

Then in any model $\mathscr{M}$,

$$\mathscr{M}(\hat{\Upsilon} z_o) = \mathbf{1} \quad \text{iff} \quad \mathbf{1} > z_0 > \mathbf{0}.$$

The $\hat{\Upsilon}$ corresponds to the I-operator above. Then a property is vague if it has typical positive, negative, and also borderline cases. Thus, for example, $A_{o\alpha}$ is vague if there are elements x_α, for which $\vdash \boldsymbol{\Delta}(A\, x_\alpha)$, elements y_α, for which $\vdash \boldsymbol{\Delta}\neg(A\, y_\alpha)$ and elements z_α, for which $\vdash \hat{\Upsilon}(A\, z_\alpha)$.

Sorites paradox

Properties should be characterized using possible worlds. Namely, they lead to different truth values in different possible worlds but remain unchanged with respect to all of them. Thus, a property is in general characterized by an intension which is a function from the set of all possible worlds into a set of extensions.

A specific property is that of "being a heap". Instead of "possible world" we will speak about *context*. The context for heaps can be modeled as a function

$$w : [0,1] \longrightarrow M_\epsilon, \qquad \begin{aligned} w(0) &= v_L, && \text{(left bound)} \\ w(0.5) &= v_S, && \text{(central point)} \\ w(1) &= v_R, && \text{(right bound)} \end{aligned}$$

for some set M_α of elements. The points v_L, v_R, v_S represent a left and right bounds, and a middle point respectively. The situation is graphically depicted in Figure 2. The context is interpretation of a formula $w_{\epsilon o}$ of FTT. We can also introduce $x \in w$ as a short for the formula $(\exists t_o)\boldsymbol{\Delta}(x \equiv w t_o)$

Using this concept, we can solve the sorites (heap) paradox[5] more realistic. Let $\mathbb{FN} \in \mathit{Form}_{(o\alpha)(\alpha o)}$ be a formula characterizing heaps. Namely, the

[5] One grain does not make a heap. Adding one grain to what is not yet a heap does not make a heap. Consequently, there are no heaps.

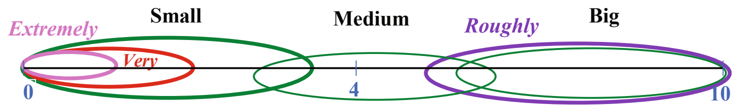


Fig. 2. Graphical representation of a possible context for heaps where $v_L = 0$, $v_S = 4$, $v_R = 10$. In the left part lay all small values, all medium values are distributed around the middle point, and all big (large) values lay in the right part of the scale.

formula $\mathbb{FN}\,wn$ expresses "n stones in a context $w \in Form_{\alpha o}$ do not form a heap". The following theorem addresses the sorites paradox and demonstrates that it is fully compatible with FTT and does not lead to contradiction.

Theorem 4. *(a)* $\vdash (\forall w)(\mathbb{FN}\,w\,0)$
(0 stones do not form a heap in any context)
(b) $\vdash (\forall w)(\forall n)(n \in w\,\mathbin{\&}\,\boldsymbol{\Delta}(w\,0.5 \leq n) \Rightarrow \neg\mathbb{FN}\,w\,n)$
(whatever number $n \geq w\,0.5$ of stones forms a heap)
(c) $\vdash (\forall w)(\exists m)(m \in w\,\mathbin{\&}\,0 < m\,\mathbin{\&}\,\hat{\Upsilon}(\mathbb{FN}\,w\,m))$
(there is a borderline number m of stones "partially" forming a heap)
(d) $\vdash (\forall w)\neg(\exists n)(n \in w\,\mathbin{\&}\,\boldsymbol{\Delta}\mathbb{FN}\,w\,n\,\mathbin{\&}\,\boldsymbol{\Delta}\neg\mathbb{FN}\,w\,(n+1))$
(in any context, there is no number n of stones surely not forming a heap such that $n+1$ surely forms a heap)
(e) $\vdash (\forall w)(\forall n)(n \in w \Rightarrow (\mathbb{FN}\,wn \Rightarrow \cdot(n \approx_w n+1) \Rightarrow \mathbb{FN}\,w(n+1)))$
(if n of stones does not form a heap then it is almost true that $n+1$ also does not form it)

A *syntactical* proof of this theorem can be found in [22].

3.4 Virtue 4: There Are Practical Extensions of FTT That Can be Effectively Implemented

Formal theory of the meaning of evaluative linguistic expressions

The considered extensions of FTT are special formal theories, the models of which constitute models of the meaning of some special linguistic expressions. The most important among them is a formal theory of the, so called, evaluative linguistic expressions. These are expressions which occur very frequently in natural language. Examples of them are the expressions *small, medium, big, twenty five, roughly one hundred, very short, more or less strong, not very tall, about twenty five, roughly small or medium, very roughly strong, weight is small, pressure is very high, extremely rich person.* Their meaning is a fundamental bearer of the vagueness phenomenon and a consequence of the indiscernibility between objects where the indiscernibility is modeled by a fuzzy equality.

There is a formal theory of the meaning of evaluative expressions developed in [22]. The meaning is characterized using a special fuzzy equality "$\sim$". The theory has altogether 11 special axioms, for example

(EV7) $\boldsymbol{\Delta}((t \Rightarrow u) \,\&\, (u \Rightarrow z)) \Rightarrow \cdot t \sim z \Rightarrow t \sim u$,
(EV8) $t \equiv t' \,\&\, z \equiv z' \Rightarrow \cdot t \sim z \Rightarrow t' \sim z'$

($t, t', u, z, z' \in \mathit{Form}_o$). The linguistic expressions with full meaning can be, in general, taken as names of intensions which are functions from the set of all possible worlds (in case of evaluative expressions, it is the set of all contexts) into a set of extensions. Intensions of evaluative expressions which are functions $W \longrightarrow \mathscr{F}(w([0,1]))$[6] are schematically depicted in Figure 3. It is relatively easy to express both intensions and extensions using formal

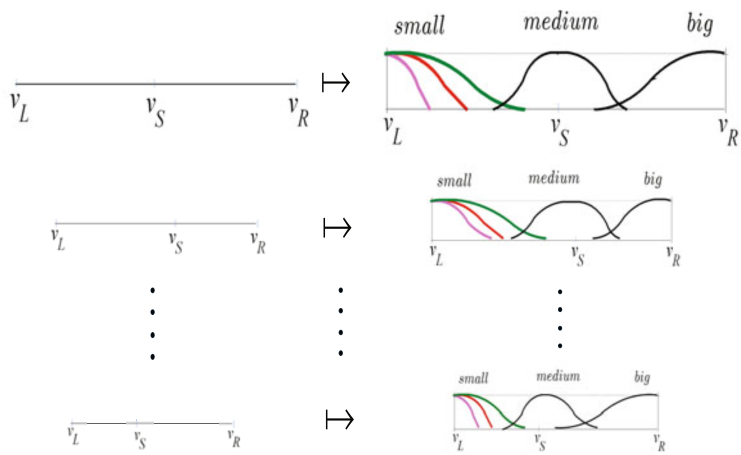


Fig. 3. Scheme of intensions of some evaluative expressions "very small, small, roughly small, medium, big"

means of FTT. The concrete models can also be constructed. Note that the latter can be then applied in straightforward way and so, this theory becomes very practical.

Theorem 5. *The theory of evaluative linguistic expressions is consistent.*

It is also possible to show that this theory addresses the sorites paradox with respect to the meaning of the evaluative expressions. For example, when replacing "heap" by "very small heap", etc., the paradox remains classically the same but, again, it is fully compatible with our theory and does not lead to contradiction. It is also formally prove that in each context there is no last

[6] $\mathscr{F}(U)$ is a set of all fuzzy sets on U.

surely small x and no first surely big x which is a typical feature of vague properties. This statement can formally be expressed as follows.

Theorem 6

$$T^{Ev} \vdash (\forall w)\neg(\exists x)(\forall y)(\boldsymbol{\Delta}(Sm\,\boldsymbol{\nu})wx\,\&(x <_w y \Rightarrow \boldsymbol{\Delta}\neg(Sm\,\boldsymbol{\nu})wy)),$$
$$T^{Ev} \vdash (\forall w)\neg(\exists x)(\forall y)(\boldsymbol{\Delta}(Bi\,\boldsymbol{\nu})wx\,\&(y <_w x \Rightarrow \boldsymbol{\Delta}\neg(Bi\,\boldsymbol{\nu})wy)).$$

Formal theory of fuzzy IF-THEN rules and linguistic descriptions

The theory of fuzzy IF-THEN rules belongs to the most successful theories developed in fuzzy logic and it has a lot of practical applications. Recall that the former are rules of the form

$$\mathsf{IF}\ X \text{ is } \mathscr{A}\ \mathsf{THEN}\ Y \text{ is } \mathscr{B}, \tag{3}$$

where 'X is $\mathscr{A}$', 'Y is $\mathscr{B}$' are evaluative predications. A (finite) set of such rules is called *linguistic description*. We can distinguish two basic approaches to the way how linguistic descriptions can be construed: relational and logical/linguistic.

The *relational approach* assumes some chosen formal system of predicate fuzzy logic. Then, certain first-order formulas $A(x), B(x)$ are assigned to the evaluative predications 'X is $\mathscr{A}$', 'Y is $\mathscr{B}$', respectively, and interpreted in a suitable formal model. Although the surface form of the rules (3) is linguistic, they are not treated in this way. The whole linguistic description is construed as a fuzzy relation resulting from the interpretation of one of two normal forms: disjunctive and conjunctive (see, e.g., the book [8], and many other publications). This approach has been well-elaborated inside predicate BL-fuzzy logic in [9, Chapter 7] and [4, 29, 30, 31], and also inside fuzzy logic with evaluated syntax in [28, Chapters 5,6]. Let us emphasize that this way of interpretation of fuzzy IF-THEN rules was developed for the approximation of functions rather than as a model of human reasoning.

Observe, however, that rules (3) can be taken as sentences of natural language. Hence, by the *logical/linguistic* approach, rules (3) are construed as genuine conditional clauses of natural language and the linguistic description is taken as a text characterizing some situation, strategy of behavior, control of some process, etc. The goal is to mimic the way how people understand natural language. Using FTT, we can construct a formal theory which includes the theory T^{Ev} of evaluative expressions so that the intension of each rule (3) can be constructed:

$$\mathrm{Int}(\mathscr{R}) := \lambda w\,\lambda w' \cdot \lambda x\,\lambda y \cdot Ev^{A}\,wx \Rightarrow Ev^{C}\,w'y \tag{4}$$

where w, w' are contexts of the antecedent and consequent of (3), respectively, Ev^A is intension of the antecedent and Ev^C intension of the consequent. The linguistic description is interpreted as a set of intensions (4) (see [21,

[25] for the details). When considering a suitable model, we obtain a formal interpretation of (4) as a function that assigns to each pair of contexts $w, w' \in W$ a fuzzy relation among objects. It is important to realize that in this case, we introduce a consistent model of the context and provide a general rule how extension can be constructed in every context.

This way of interpretation requires a special inference procedure called *perception-based logical deduction* (see [19, 27]). The main idea is to consider the linguistic description as a specific text, which has a *topic* (what we are speaking about) and *focus* (what is the new information — for the detailed linguistic analysis of these concepts, see, e.g., [10]). Each rule is understood as vague local information about existing relation between X and Y. The procedure can distinguish among the rules and if some observation is given, to derive a conclusion which is in accordance with the human way of reasoning.

Formal theory of intermediate quantifiers

Generalized quantifiers occur quite often in natural language. Recall that these are words such as *most, a lot of, many, a few, a great deal of, a large part of, etc.*. A general theory of generalized quantifiers was initiated by A. Mostowski in [15] and further elaborated by P. Lindström, D. Westerståhl, E. L. Keenan, J. Barwise, R. Cooper ([13, 14, 34, 32]). Generalized quantifiers were introduced into fuzzy logic by L. A. Zadeh [35] and further elaborated by P. Hájek [9, Chapter 8], I. Glöckner [7], M. Holčapek and A. Dvořák [5, 11] and by a few other people.

The class of intermediate quantifiers has been introduced and analyzed in detail in [33]. Examples of them are *many, a lot of, most, almost all*, etc. The main idea of how their semantics can be captured is the following: Intermediate quantifiers refer to elements taken from a class that is "smaller" than the original universe in a specific way. Namely, they are classical quantifiers "for all" or "exists" taken over a class of elements that is determined using an appropriate evaluative expression. Classical logic has no substantiation for why and how the range of quantification should be made smaller. In fuzzy logic, we can apply the theory of evaluative linguistic expressions as follows (for the details, see [23]). Let the formula Ev represent an evaluative predication and μ a formula representing measure of fuzzy sets. Then we define

$$(Q^{\forall}_{Ev}\, x_\alpha)(B_{o\alpha}, A_{o\alpha}) := (\exists z_{o\alpha})((\boldsymbol{\Delta}(z_{o\alpha} \subseteq B_{o\alpha})\, \&\, (\forall x_\alpha)(z_{o\alpha}\, x_\alpha \Rightarrow A_{o\alpha} x_\alpha)) \wedge Ev((\mu B_{o\alpha}) z_{o\alpha})), \tag{5}$$

$$(Q^{\exists}_{Ev}\, x_\alpha)(B_{o\alpha}, A_{o\alpha}) := (\exists z_{o\alpha})((\boldsymbol{\Delta}(z_{o\alpha} \subseteq B_{o\alpha})\, \&\, (\exists x_\alpha)(z_{o\alpha}\, x_\alpha \wedge A_{o\alpha} x_\alpha)) \wedge Ev((\mu B_{o\alpha}) z_{o\alpha})). \tag{6}$$

Interpretation of (5) is the following: there is a fuzzy set $z_{o\alpha}$ of objects having the property $B_{o\alpha}$, the size of which (determined by the measure μ) is characterized by the evaluative expression Ev, and all these objects also have the property $A_{o\alpha}$. The interpretation of (6) is similar. The property is here, for simplicity, represented by a fuzzy set. However, it is also possible to introduce possible worlds. Note that these quantifiers belong to the wide class of generalized quantifiers [14, 32, 34], which in fuzzy logic have been introduced also in [5, 7, 9, 11, 26].

The theory also includes classical quantifiers. Furthermore, special natural language quantifiers can be specified, e.g., by the following formulas:

$$\begin{aligned} \mathit{Most} &:= Q^{\forall}_{\mathit{Very\ big}} & \mathit{Many} &:= Q^{\forall}_{\mathit{Big}} \\ \mathit{Several} &:= Q^{\forall}_{\mathit{Small}} & \mathit{Some} &:= Q^{\exists}_{\mathit{Small}} \end{aligned}$$

In [33], altogether total of 105 *generalized syllogisms* were informally introduced (including the basic Aristotelian syllogisms). All of them are valid also in this theory (see [16]), for example:

ATK-I:

$$\begin{array}{l} \text{All } M \text{ are } Y \\ \text{Most } X \text{ are } M \\ \hline \text{Many } X \text{ are } Y \end{array}$$

PKI-III:

$$\begin{array}{l} \text{Almost all } M \text{ are } Y \\ \text{Many } M \text{ are } X \\ \hline \text{Some } X \text{ are } Y \end{array}$$

4 Conclusion

This paper addresses the fuzzy type theory which is the higher-order fuzzy logic. There are several kinds of this logic depending on the chosen structure of truth values. The higher-order (fuzzy) logic is still not fully appreciated despite its high explicative power. Our goal is to point out several important virtues of it to convince the reader that this logic is worth of studying and also using.

The main discussed virtues were the following: (1) FTT has a simple and highly uniform syntax; (2) Semantics of FTT is based on a small collection of well-established ideas; (3) FTT is a highly expressive logic; (4) There are practical extensions of FTT that can be effectively implemented.

References

1. Andrews, P.: An Introduction to Mathematical Logic and Type Theory: To Truth Through Proof. Kluwer, Dordrecht (2002)
2. Běhounek, L., Cintula, P.: From fuzzy logic to fuzzy mathematics: A methodological manifesto. Fuzzy Sets and Systems 157(5), 642–646 (2006)
3. Church, A.: A formulation of the simple theory of types. J. Symb. Logic 5, 56–68 (1940)
4. Daňková, M.: Representation of logic formulas by normal forms. Kybernetika 38, 711–728 (2002)
5. Dvořák, A., Holčapek, M.: L-fuzzy quantifiers of the type ⟨1⟩ determined by measures. Fuzzy Sets and Systems 160, 3425–3452 (2009)

6. Farmer, W.M.: The seven virtues of simple type theory. Journal of Applied Logic 6, 267–286 (2008)
7. Glöckner, I.: Fuzzy Quantifiers: A Computational Theory. Springer, Berlin (2006)
8. Gottwald, S.: Fuzzy Sets and Fuzzy Logic. The Foundations of Application – from a Mathematical Point of View, Vieweg. Braunschweig/Wiesbaden and Teknea, Toulouse (1993)
9. Hájek, P.: Metamathematics of Fuzzy Logic. Kluwer, Dordrecht (1998)
10. Hajičová, E., Partee, B.H., Sgall, P.: Topic-Focus Articulation, Tripartite Structures, and Semantics Content. Kluwer, Dordrecht (1998)
11. Holčapek, M.: Monadic L-fuzzy quantifiers of the type $\langle 1^n, 1 \rangle$. Fuzzy Sets and Systems 159, 1811–1835 (2008)
12. Keefe, R.: Theories of Vagueness. Cambridge University Press, Cambridge (2000)
13. Keenan, E., Westerståhl, D.: Quantifiers in formal and natural languages. In: van Benthem, J., ter Meulen, A. (eds.) Handbook of Logic and Language, pp. 837–893. Elsevier, Amsterdam (1997)
14. Lindström, P.: First order predicate logic with generalized quantifiers. Theoria 32, 186–195 (1966)
15. Mostowski, A.: On a generalization of quantifiers. Fundamenta Mathematicae 44, 12–36 (1957)
16. Murinová, P., Novák, V.: A formal theory of generalized intermediate syllogisms. Fuzzy Sets and Systems (2010)
17. Novák, V.: Fuzzy type theory as higher order fuzzy logic. In: Proc. 6th Int. Conference on Intelligent Technologies (InTech 2005), December 14-16. Fac. of Science and Technology, Assumption University, Bangkok (2005)
18. Novák, V.: On fuzzy type theory. Fuzzy Sets and Systems 149, 235–273 (2005)
19. Novák, V.: Perception-based logical deduction. In: Reusch, B. (ed.) Computational Intelligence, Theory and Applications, pp. 237–250. Springer, Berlin (2005)
20. Novák, V.: Fuzzy sets as a special mathematical model of vagueness phenomenon. In: Reusch, B. (ed.) Computational Intelligence, Theory and Applications, pp. 683–690. Springer, Heidelberg (2006)
21. Novák, V.: Mathematical fuzzy logic in modeling of natural language semantics. In: Wang, P., Ruan, D., Kerre, E. (eds.) Fuzzy Logic – A Spectrum of Theoretical & Practical Issues, pp. 145–182. Elsevier, Berlin (2007)
22. Novák, V.: A comprehensive theory of trichotomous evaluative linguistic expressions. Fuzzy Sets and Systems 159(22), 2939–2969 (2008)
23. Novák, V.: A formal theory of intermediate quantifiers. Fuzzy Sets and Systems 159(10), 1229–1246 (2008)
24. Novák, V.: EQ-algebra-based fuzzy type theory and its extensions. Logic Journal of the IGPL (to appear, 2010)
25. Novák, V., Lehmke, S.: Logical structure of fuzzy IF-THEN rules. Fuzzy Sets and Systems 157, 2003–2029 (2006)
26. Novák, V., Perfilieva, I. (eds.): Discovering the World With Fuzzy Logic. Studies in Fuzziness and Soft Computing, vol. 57. Springer, Heidelberg (2000)
27. Novák, V., Perfilieva, I.: On the semantics of perception-based fuzzy logic deduction. International Journal of Intelligent Systems 19, 1007–1031 (2004)
28. Novák, V., Perfilieva, I., Močkoř, J.: Mathematical Principles of Fuzzy Logic. Kluwer, Boston (1999)

29. Perfilieva, I.: Fuzzy logic normal forms for control law representation. In: Verbruggen, H., Zimmermann, H.-J., Reusch, B. (eds.) Fuzzy Algorithms for Control, pp. 111–125. Kluwer, Boston (1999)
30. Perfilieva, I.: Logical approximation. Soft Computing 7(2), 73–78 (2002)
31. Perfilieva, I.: Normal forms in BL-algebra of functions and their contribution to universal approximation. Fuzzy Sets and Systems 143, 111–127 (2004)
32. Peters, S., Westerståhl, D.: Quantifiers in Language and Logic. Claredon Press, Oxford (2006)
33. Peterson, P.: Intermediate Quantifiers. Logic, linguistics, and Aristotelian semantics. Ashgate, Aldershot (2000)
34. Westerståhl, D.: Quantifiers in formal and natural languages. In: Gabbay, D., Guenthner, F. (eds.) Handbook of Philosophical Logic, vol. IV, pp. 1–131. D. Reidel, Dordrecht (1989)
35. Zadeh, L.A.: A computational approach to fuzzy quantifiers in natural languages. Computers and Mathematics 9, 149–184 (1983)

Semilinear Space, Galois Connections and Fuzzy Relation Equations

Irina Perfilieva

Institute for Research and Applications of Fuzzy Modeling,
University of Ostrava 30. dubna 22, 701 03 Ostrava 1, Czech Republic
e-mail: Irina.Perfilieva@osu.cz

Abstract. We introduce a notion of an idempotent semilinear space and consider two systems of linear-like equations. These systems are equivalent to systems of fuzzy relation equations with sup-$*$ and inf-$\rightarrow$ compositions. Moreover, because the two types of systems of linear-like equations are dual according to this theory, it is sufficient to investigate only one system.

Keywords: Semilinear space, Residuated lattice, System of fuzzy relation equations, Fixed point.

1 Introduction

The aim of this paper is twofold: first, to aid in the formalization and unification of tools and methods used in the theory of fuzzy relation equations, and second, to propose a generalization of the theory of linear spaces. As is known from the extant literature, there are at least two types of systems of fuzzy relation equations that differ in types of composition [1, 2, 3, 4]. However, results about the solvability and the structure of solution sets for both types of composition are, in some sense, dual. Additionally, there is a profound theory of linear spaces wherein the problem of determining the solvability of systems of linear equations is entirely solved. Thus, our motivation was to find a proper generalization of the theory of linear spaces that can also serve as a theoretical platform for the analysis of systems of fuzzy relation equations.

In this paper, we will show that the theory of Galois connections can be successfully used in explaining the above mentioned duality and characterizing the solvability. In more details, if solvability is connected with a characterization of the vectors on the right-hand sides, then there exists a Galois connection between a set of admissible right-hand sides and a set of solutions. Moreover, on the basis of this theory, two types of systems of linear-like equations are dual, and thus it suffices to study only one of them.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 71–79.
springerlink.com

2 Idempotent Semilinear Spaces

We recall that a linear (vector) space is a special case of a module over a ring, i.e., a linear space is a unitary module over a field [5]. In this paper, we will be dealing with a unitary semimodule over a commutative semiring [6, 7], which will be called a *semilinear space*. Moreover, our semilinear space will be an idempotent structure with respect to its main operation.

Definition 1. *Let $\mathcal{R} = (R, +, \cdot, 0, 1)$ be a commutative semiring and $\mathcal{V} = (V, +, \bar{0})$ a commutative monoid. We say that $\mathcal{V}$ is a (left) semilinear space over $\mathcal{R}$ if an external (left) multiplication $\lambda : \bar{x} \mapsto \lambda\bar{x}$ is defined, where $\lambda \in R$ and $\bar{x} \in V$. Moreover, the following mutual properties are fulfilled for all $\bar{x}, \bar{y} \in V$ and $\lambda, \mu \in R$:*

(SLS1) $\lambda(\bar{x} + \bar{y}) = \lambda\bar{x} + \lambda\bar{y}$,
(SLS2) $(\lambda + \mu)\bar{x} = \lambda\bar{x} + \mu\bar{x}$,
(SLS3) $(\lambda \cdot \mu)\bar{x} = \lambda(\mu\bar{x})$,
(SLS4) $1\bar{x} = \bar{x}$,
(SLS5) $\lambda\bar{0} = \bar{0}$.

When $\mathcal{R}$ is clear from the context, we will shorten "left semilinear space over $\mathcal{R}$" to "semilinear space." Elements of a semilinear space will be distinguished by an overline.

Example 1. Let $\mathcal{R} = (R, +, \cdot, 0, 1)$ be a commutative semiring. Let R^n $(n \geq 1)$ be the set of n-dimensional vectors whose components are elements of R, i.e. $R^n = \{\bar{x} = (x_1, \ldots, x_n) \mid x_1 \in R, \ldots, x_n \in R\}$. Let $\bar{0} = (0, \ldots, 0)$ and

$$\bar{x} + \bar{y} = (x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n).$$

Then, $\mathcal{R}^n = (R^n, +, \bar{0})$ is a commutative monoid. For any $\lambda \in R$, external multiplication $\lambda\bar{x}$ is defined by

$$\lambda\bar{x} = \lambda(x_1, \ldots, x_n) = (\lambda \cdot x_1, \ldots, \lambda \cdot x_n).$$

Then, $\mathcal{R}^n$ is a semilinear space over $\mathcal{R}$.

Semilinear space $\mathcal{R}^n$, $n \geq 1$, (see Example 1) will be called *vectorial semilinear space* over $\mathcal{R}$.

Definition 2. *Semilinear space $\mathcal{V}$ over $\mathcal{R}$ is called idempotent if the operations $+$ in both $\mathcal{V}$ and $\mathcal{R}$ are idempotent.*

Let $\mathcal{V} = (V, +, \bar{0})$ be an idempotent semilinear space. Then

$$\bar{x} \leq \bar{y} \iff \bar{x} + \bar{y} = \bar{y}, \tag{1}$$

is the *natural order* on $\mathcal{V}$. Therefore, $(V, \leq)$ is a bounded $\vee$-semilattice where $\bar{x} \vee \bar{y} = \bar{x} + \bar{y} = \sup\{\bar{x}, \bar{y}\}$, and $\bar{0}$ is a bottom element.

It may happen (see Example 2 below) that two idempotent semilinear spaces $\mathcal{V}_1 = (V, +_1, \bar{0}_1)$ and $\mathcal{V}_2 = (V, +_2, \bar{0}_2)$ with the same support V determine *dual* (or reverse) natural orders $\leq_1$ and $\leq_2$ on V, i.e.,

$$\bar{x} \leq_1 \bar{y} \iff \bar{y} \leq_2 \bar{x}.$$

In this case, $\leq_2$ is simply denoted $\geq_1$ or $\leq_1^d$. With respect to $\leq_1$, $(\mathcal{V}_2, \geq_1)$ is a $\wedge$-semilattice where $\bar{x} \wedge \bar{y} = \bar{x} +_2 \bar{y} = \inf\{\bar{x}, \bar{y}\}$. We will call $\mathcal{V}_1$ a $\vee$-*semilinear space*, and $\mathcal{V}_2$ a $\wedge$-*semilinear space*. Moreover, if $\mathcal{V}_1$ and $\mathcal{V}_2$ are idempotent semilinear spaces over the same semiring, then we will call them *dual*. It is easy to see that for dual semilinear spaces, the *Principle of Duality* for ordered sets holds true.

Example 2. Let $\mathcal{L} = (L, \vee, \wedge, *, \rightarrow, 0, 1)$ be an integral, residuated, commutative l-monoid and $\mathcal{L}_\vee = (L, \vee, *, 0, 1)$ a commutative $\vee$-semiring. L^n $(n \geq 1)$ is a set of n-dimensional vectors as in Example 1.

1. $\mathcal{L}_\vee^n = (L^n, \vee, \bar{0})$ is an idempotent commutative monoid, where $\bar{0} = (0, \ldots, 0) \in L^n$, and for any $\bar{x}, \bar{y} \in L^n$,

$$\bar{x} \vee \bar{y} = (x_1, \ldots, x_n) \vee (y_1, \ldots, y_n) = (x_1 \vee y_1, \ldots, x_n \vee y_n).$$

   The order on $\mathcal{L}_\vee^n$ is determined by $\vee$ so that $\bar{x} \leq \bar{y}$ if and only if $x_1 \leq y_1, \ldots, x_n \leq y_n$. For any $\lambda \in L$, external multiplication $\lambda\bar{x}$ is defined by

$$\lambda\bar{x} = \lambda(x_1, \ldots, x_n) = (\lambda * x_1, \ldots, \lambda * x_n).$$

   $\mathcal{L}_\vee^n$ with external multiplication $\lambda : \bar{x} \mapsto \lambda\bar{x}$ is an (idempotent) $\vee$-semilinear space over $\mathcal{L}_\vee$.
2. $\mathcal{L}_\wedge^n = (L^n, \wedge, \bar{1})$ is an idempotent commutative monoid where $\bar{1} = (1, \ldots, 1) \in L^n$, and for any $\bar{x}, \bar{y} \in L^n$,

$$(x_1, \ldots, x_n) \wedge (y_1, \ldots, y_n) = (x_1 \wedge y_1, \ldots, x_n \wedge y_n).$$

   The natural order on $\mathcal{L}_\wedge^n$ is determined by $\wedge$, and this ordering is dual to $\leq$, which was introduced on L^n in case 1 above. We will denote the natural order on $\mathcal{L}_\wedge^n$ by $\leq^d$, so that $\bar{x} \leq^d \bar{y}$ if and only if $\bar{x} \geq \bar{y}$. Alternatively, $\bar{x} \leq^d \bar{y}$ if and only if $x_1 \geq y_1, \ldots, x_n \geq y_n$. For any $\lambda \in L$, let us define external multiplication $\lambda \diagdown \bar{x}$ by

$$\lambda \diagdown (x_1, \ldots, x_n) = (\lambda \rightarrow x_1, \ldots, \lambda \rightarrow x_n).$$

   $\mathcal{L}_\wedge^n$ with the external multiplication $\lambda : \bar{x} \mapsto \lambda \diagdown \bar{x}$ is an (idempotent) $\wedge$-semilinear space over $\mathcal{L}_\vee$.

$\vee$-semilinear space $\mathcal{L}_\vee^n$ and $\wedge$-semilinear space $\mathcal{L}_\wedge^n$ are dual.

2.1 Galois Connections in Semilinear Spaces

Let us recall that a Galois connection between two ordered sets $(A, \leq)$ and $(B, \leq)$ is a pair (h, g) of antitone mappings $h : A \to B$ and $g : B \to A$ such that $h \circ g \geq \mathrm{id_A}$ and $g \circ h \geq \mathrm{id_B}$ ($\circ$ denotes the composition of two mappings so that, e.g., for all $x \in A$, $(h \circ g)(x) = g(h(x))$).

In this section, we will show that two dual idempotent semilinear spaces can be connected by various Galois connections.

Theorem 1. *Let $\mathcal{L}_\vee^m$, $m \geq 1$, be a $\vee$-semilinear space and $\mathcal{L}_\wedge^n$, $n \geq 1$, a $\wedge$-semilinear space, both over $\mathcal{L}_\vee$ (see Example 2). Let $\leq$ and $\leq^d$ be natural orders on $\mathcal{L}_\vee^m$ and $\mathcal{L}_\wedge^n$ respectively. Then*

(i) for each $\lambda \in L$, mappings $\bar{x} \mapsto \lambda\bar{x}$ and $\bar{y} \mapsto \lambda \diagdown \bar{y}$ establish a Galois connection between $(\mathcal{L}_\vee^n, \leq)$ and $(\mathcal{L}_\wedge^n, \leq^d)$,

(ii) for each $n \times m$ matrix $A \in L^{n\times m}$ with transpose A^*, mappings $h_A : L^m \to L^n$ and $g_{A^*} : L^n \to L^m$ given by

$$h_A(\bar{x})_i = a_{i1} * x_1 \vee \cdots \vee a_{im} * x_m, \quad i = 1, \ldots, n, \tag{2}$$

and

$$g_{A^*}(\bar{y})_j = (a_{1j} \to y_1) \wedge \cdots \wedge (a_{nj} \to y_n), \quad j = 1, \ldots, m, \tag{3}$$

establish a Galois connection between $(\mathcal{L}_\vee^m, \leq)$ and $(\mathcal{L}_\wedge^n, \leq^d)$.

Throughout this paper, let $\mathcal{L} = \langle L, \vee, \wedge, *, \to, 0, 1\rangle$ be an integral, residuated, commutative l-monoid (*a residuated lattice*), U a non-empty set and L^U a set of L-valued functions on U. Fuzzy subsets of U are identified with L-valued functions on U (membership functions).

3 Systems of Fuzzy Relation Equations and Their Semilinear Analogs

Let U and V be two universes (not necessary different), $A_i \in L^U$, $B_i \in L^V$ arbitrarily chosen fuzzy subsets of respective universes, and $R \in L^{U\times V}$ a fuzzy subset of $U \times V$. This last item is called a fuzzy relation. Lattice operations $\vee$ and $\wedge$ are used for the union and intersection of fuzzy sets, respectively. Two other binary operations $*, \to$ of $\mathcal{L}$ are used for compositions—binary operations on $L^{U\times V}$. We will consider two of them: sup-$*$ -composition, usually denoted $\circ$, and inf-$\to$composition usually denoted $\lhd$. The first was introduced by L. Zadeh [8] and the second by W. Bandler and L. Kohout [9]. We will demonstrate definitions of both compositions on particular examples of set-relation compositions $A \circ R$ and $A \lhd R$, where $A \in L^U$ and $R \in L^{U\times V}$:

$$(A \circ R)(v) = \bigvee_{u\in U} (A(u) * R(u, v)),$$

$$(A \lhd R)(v) = \bigwedge_{u\in U} (A(u) \to R(u, v)).$$

Remark 1. Let us remark that both compositions can be considered as set-set compositions where R is assumed to be replaced by a fuzzy set. They are used in this reduced form later in instances of systems of fuzzy relation equations.

By a *system of fuzzy relation equations with sup-*-composition*, we mean the following system of equations

$$A_i \circ R = B_i, \text{ or } \bigvee_{u \in U} (A_i(u) * R(u,v)) = B_i(v), \quad 1 \le i \le n, \tag{4}$$

considered with respect to unknown fuzzy relation $R \in L^{U \times V}$. Its counterpart is a *system of fuzzy relation equations with* inf-$\rightarrow$*composition*

$$A_j \lhd R = D_j, \text{ or } \bigwedge_{u \in U} (A_j(u) \rightarrow R(u,v)) = D_j(v), \quad 1 \le j \le m, \tag{5}$$

also considered with respect to unknown $R \in L^{U \times V}$. System (4) and its potential solutions are well investigated in the literature (see e.g. [1, 2, 10, 11, 12, 13, 14, 15]). On the other hand, investigations of the solvability of (5) are not so intensive (see [2, 4, 16]).

4 Systems of Equations in Semilinear Spaces $\mathcal{L}_\vee^m$ and $\mathcal{L}_\wedge^n$

4.1 System of Equations in Semilinear Space $\mathcal{L}_\vee^m$

Let $\mathcal{L}_\vee^m$, with $m \ge 1$, be a $\vee$-semilinear space over $\mathcal{L}_\vee$, and $\mathcal{L}_\wedge^n$, with $n \ge 1$, be a $\wedge$-semilinear space over $\mathcal{L}_\vee$. Let $n \times m$ matrix $A = (a_{ij})$, vector $\bar{b} = (b_1, \ldots, b_n) \in L^n$, and vector $\bar{d} = (b_1, \ldots, d_m) \in L^m$ have components from L. The systems of equations

$$\begin{gathered} a_{11} * x_1 \vee \cdots \vee a_{1m} * x_m = b_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ a_{n1} * x_1 \vee \cdots \vee a_{nm} * x_m = b_n, \end{gathered} \tag{6}$$

and

$$\begin{gathered} (a_{11} \rightarrow y_1) \wedge \cdots \wedge (a_{n1} \rightarrow y_n) = d_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ (a_{1m} \rightarrow y_1) \wedge \cdots \wedge (a_{nm} \rightarrow y_n) = d_m, \end{gathered} \tag{7}$$

are considered with respect to unknown vectors $\bar{x} = (x_1 \ldots, x_m) \in L^m$ and $\bar{y} = (y_1 \ldots, y_n) \in L^n$. By (2) and (3), systems (6) and (7) can be respectively represented as follows:

$$h_A(\bar{x}) = \bar{b}.$$

and

$$g_{A^*}(\bar{y}) = \bar{d}.$$

It is easily seen that system (6) is an instance of (4) specified by $U = \{u_1, \ldots, u_m\}$, $V = \{v_1, \ldots, v_n\}$, fixed $v \in V$, and $A_i(u_j) = a_{ij}$, $B_i(v) = b_i$, $R(u_j, v) = x_j$. Similarly, system (7) is an instance of (5) specified by $U = \{u_1, \ldots, u_n\}$, $V = \{v_1, \ldots, v_m\}$, fixed $v \in V$, and $A_i(u_j) = a_{ji}$, $D_i(v) = d_i$, $R(u_j, v) = y_j$.

4.2 Solvability in Terms of Galois Connection

Below, we will give results regarding solvability and the solutions of systems (6) and (7) represented as:

$$A\bar{x} = \bar{b}, \quad (h_A(\bar{x}) = \bar{b}),$$
$$A^*\diagdown\bar{y} = \bar{d}, \quad (g_{A^*}(\bar{y}) = \bar{d}),$$

where mappings h_A, g_{A^*} establish a Galois connection between dually ordered spaces $(\mathcal{L}^m_\vee, \leq)$ and $(\mathcal{L}^n_\wedge, \leq^d)$. Therefore, any result about the solvability of one system has its dual counterpart, which can be obtained by

- replacing h_A by g_{A^*}, and vice versa,
- replacing $\leq$ by $\geq$, and vice versa,
- replacing $\vee$ by $\wedge$, and vice versa.

For the reader's convenience, we will formulate both dual results about the solvability and the solutions of systems (6) and (7).

Theorem 2. *Let A be a given matrix, and h_A and g_{A^*} establish a Galois connection between semilinear spaces $\mathcal{L}^m_\vee$ and $\mathcal{L}^n_\wedge$. Then,*

(i) System (6) is solvable if and only if $\bar{b}$ is a closed element of $\mathcal{L}^n_\wedge$ with respect to the closure operator $g_{A^*} \circ h_A$, or if and only if

$$\bar{b} = h_A(g_{A^*}(\bar{b})) = A(A^*\diagdown\bar{b}). \tag{8}$$

(ii) System (7) is solvable if and only if $\bar{d}$ is a closed element of $\mathcal{L}^m_\vee$ with respect to the closure operator $h_A \circ g_{A^*}$, or if and only if

$$\bar{d} = g_{A^*}(h_A(\bar{d})) = A^*\diagdown(A\bar{d}). \tag{9}$$

Remark 2. By (8), the right-hand side vector $\bar{b} \in L^n$ of a solvable system (6) is a *fixed point* of the closure operator $g_{A^*} \circ h_A$ determined by the matrix of coefficients A. Similarly by (9), the right-hand side vector $\bar{d} \in L^m$ of a solvable system (7) is a fixed point of the closure operator $h_A \circ g_{A^*}$.

Remark 3. By Theorem 1, $A(A^*\diagdown\bar{y}) \leq \bar{y}$ so that the operator $g_{A^*} \circ h_A$ is a closure in $\mathcal{L}^n_\wedge$ ordered by the dual ordering $\leq^d$. In general, a closure operator in a dually ordered space is called an opening operator with respect to the reverse, i.e., genuine, ordering $\leq$. We will not, however, use this term.

Corollary 1. *Let the conditions of Theorem 2 be fulfilled. Then,*

(i) $\bar{b}$ is a fixed point of $g_{A^*} \circ h_A$ if and only if there exists $\bar{x} \in L^m$ such that $h_A(\bar{x}) = \bar{b}$, or $A\bar{x} = \bar{b}$.
(ii) $\bar{d}$ is a fixed point of $h_A \circ g_{A^*}$ if and only if there exists $\bar{y} \in L^n$ such that $g_{A^*}(\bar{y}) = \bar{d}$, or $A^*\diagdown\bar{y} = \bar{d}$.

Corollary 2. *Let the conditions of Theorem 2 be fulfilled. Then,*

(i) for each $\bar{x} \in L^m$, $A(A^*\diagdown A\bar{x}) = A\bar{x}$,
(ii) for each $\bar{y} \in L^n$, $A^*\diagdown A(A^*\diagdown\bar{y}) = A^*\diagdown\bar{y}$.

Theorem 3. *Let A be a given matrix, $g_{A^*} \circ h_A$ a closure operator on $\mathcal{L}^n_\wedge$, $h_A \circ g_{A^*}$ a closure operator on $\mathcal{L}^m_\vee$. Then,*

(i) the set $cl^*_A(L^n)$ of fixed points of $g_{A^*} \circ h_A$ is a semilinear subspace of $\mathcal{L}^n_\vee$.
(ii) the set $cl_A(L^m)$ of fixed points of $h_A \circ g_{A^*}$ is a semilinear subspace of $\mathcal{L}^m_\wedge$.

Theorem 4. *Let systems (6) and (7) be specified by $n \times m$ matrix A and vectors $\bar{b} \in L^n$, $\bar{d} \in L^m$, respectively. Then,*

(i) if $\bar{b}$ is a fixed point of $g_{A^*} \circ h_A$, then $g_{A^*}(\bar{b}) = A^*\diagdown\bar{b}$ is a solution of system (6),
(ii) if $\bar{d}$ is a fixed point of $h_A \circ g_{A^*}$, then $h_A(\bar{d}) = A\bar{d}$ is a solution of system (7).

Theorem 5. *Let systems (6) and (7) be specified by $n \times m$ matrix A and vectors $\bar{b} \in L^n$, $\bar{d} \in L^m$, respectively. Then,*

(i) h_A restricted to the set of fixed points $cl_A(L^m)$ is a bijection between $cl_A(L^m)$ and $cl^*_A(L^n)$,
(ii) g_{A^*} restricted to the set of fixed points $cl^*_A(L^n)$ is a bijection between $cl^*_A(L^n)$ and $cl_A(L^m)$,
(iii) restriction $g_{A^*}|_{cl^*_A(L^n)}$ is inverse of the restriction $h_A|_{cl_A(L^m)}$.

Let $\equiv_{h_A}$ be an equivalence relation on $\mathcal{L}^m_\vee$ such that

$$\bar{x}_1 \equiv_{h_A} \bar{x}_2 \iff A(\bar{x}_1) = A(\bar{x}_2).$$

Similarly, let $\equiv_{g_{A^*}}$ be an equivalence relation on $\mathcal{L}^n_\wedge$ such that

$$\bar{y}_1 \equiv_{g_{A^*}} \bar{y}_2 \iff A^*\diagdown\bar{y}_1 = A^*\diagdown\bar{y}_2.$$

Define $[\bar{x}]_{\equiv_{h_A}}$ as an equivalence class of $\bar{x}$ with respect to $\equiv_{h_A}$, and $[\bar{y}]_{\equiv_{g_{A^*}}}$ an equivalence class of $\bar{y}$ with respect to $\equiv_{g_{A^*}}$.

Lemma 1. *Let A be a $n \times m$ matrix. Then,*

(i) For all $\bar{x} \in L^m$, $[\bar{x}]_{\equiv_{h_A}} = [A^* \diagdown A\bar{x}]_{\equiv_{h_A}}$, where $A^* \diagdown A\bar{x} \in cl_A(L^m)$ is a fixed point of $h_A \circ g_{A^*}$.

(ii) For all $\bar{y} \in L^n$, $[\bar{y}]_{\equiv_{g_{A^*}}} = [A(A^* \diagdown \bar{y})]_{\equiv_{g_{A^*}}}$, where $A(A^* \diagdown \bar{y}) \in cl^*_A(L^n)$ is a fixed point of $g_{A^*} \circ h_A$.

Theorem 6. *Let systems (6) and (7) be specified by $n \times m$ matrix A and vectors $\bar{b} \in L^n$, $\bar{d} \in L^m$, respectively. Then,*

(i) $[A^* \diagdown \bar{b}]_{\equiv_{h_A}}$ is a set of solutions of (6) with the righthand side given by $\bar{b}$, i.e.,

$$\bar{x} \in [A^* \diagdown \bar{b}]_{\equiv_{h_A}} \Leftrightarrow A\bar{x} = \bar{b}.$$

Moreover, $A^* \diagdown \bar{b} \in cl_A(L^m)$, and $A^* \diagdown \bar{b}$ is the greatest element in $[A^* \diagdown \bar{b}]$.

(ii) $[A\bar{d}]_{\equiv_{g_{A^*}}}$ is a set of solutions of (7) with the righthand side given by $\bar{d}$, i.e.

$$\bar{y} \in [A\bar{d}]_{\equiv_{g_{A^*}}} \Leftrightarrow A^* \diagdown \bar{y} = \bar{d}.$$

Moreover, $A\bar{d} \in cl^*_A(L^n)$, and $A\bar{d}$ is the least element in $[A\bar{d}]$.

5 Conclusion

In this paper, we showed that the theory of semilinear spaces and Galois connections can be successfully used in characterizing the solvability and solutions sets of systems of linear-like equations. The solvability is characterized by the relationship between vectors of right-hand sides and solutions. Moreover, because two types of systems of linear-like equations are shown to be dual on the basis of this theory, only one of them was investigated.

Acknowledgements. The research has been supported by the grant IAA108270902 of GA AVČR, and by the project MSM 6198898701 of the MŠMT ČR.

References

1. Di Nola, A., Sessa, S., Pedrycz, W., Sanchez, E.: Fuzzy Relation Equations and Their Applications to Knowledge Engineering. Kluwer, Boston (1989)
2. De Baets, B.: Analytical solution methods for fuzzy relation equations. In: Dubois, D., Prade, H. (eds.). The Handbooks of Fuzzy Sets Series, vol. 1, pp. 291–340. Kluwer, Dordrecht (2000)
3. Nosková, L., Perfilieva, I.: System of fuzzy relation equations with sup-*composition in semi-linear spaces: minimal solutions. In: Proc. FUZZ-IEEE Conf. on Fuzzy Systems, London, pp. 1520–1525 (2007)
4. Perfilieva, I., Nosková, L.: System of fuzzy relation equations with inf→composition: complete set of solutions. Fuzzy Sets and Systems 159, 2256–2271 (2008)

5. Raikov, D.A.: Vector spaces. FizMatLit, Moscow (1962)
6. Kurosh, A.G.: Lectures on General Algebra. FizMatLit, Moscow (1962)
7. Golan, J.S.: Semirings and their Applications. Kluwer Academic Pulishers, Dordrecht (1999)
8. Zadeh, L.A.: The concept of a linguistic variable and its application to approximate reasoning I, II, III. Information Sciences 8-9, 199–257, 301–357, 43–80 (1975)
9. Bandler, W., Kohout, L.: Semantics of implication operators and fuzzy relational products. Int. J. Man-Machine Studies 12, 89–116 (1980)
10. Gottwald, S.: On the existence of solutions of systems of fuzzy equations. Fuzzy Sets and Systems 12, 301–302 (1984)
11. Klawonn, F.: Fuzzy points, fuzzy relations and fuzzy functions. In: Novák, V., Perfilieva, I. (eds.) Discovering the World with Fuzzy Logic, pp. 431–453. Springer, Berlin (2000)
12. Perfilieva, I., Tonis, A.: Compatibility of systems of fuzzy relation equations. Int. J. of General Systems 29, 511–528 (2000)
13. Perfilieva, I., Gottwald, S.: Solvability and approximate solvability of fuzzy relation equations. Int. J. of General Systems 32, 361–372 (2003)
14. Sanchez, E.: Resolution of composite fuzzy relation equations. Information and Control 30, 38–48 (1976)
15. Wang, X.: Infinite fuzzy relational equations on a complete Brouwerian lattice. Fuzzy Sets and Systems 138, 657–666 (2003)
16. Štěpnička, M., Jayaram, B.: On the suitability of the bandler-kohout subproduct as an inference mechanism. IEEE Trans. on Fuzzy Syst. (2010) (to appear)

Quantitative Logic: A Quantitative Approach to Many-Valued Logics

Guo-jun Wang and Hong-jun Zhou

College of Mathematics and Information Science, Shaanxi Normal University, Xi'an 710062, China
e-mail: {gjwang,hjzhou}@snnu.edu.cn

Abstract. Quantitative logic mainly provides a quantitative approach to evaluating the degree of truth of a formula in a system of many-valued logic. Based on the fundamental notion of truth degree, one can introduce the notions of similarity degree and of pseudo-metric between formulas, and can also introduce an effective method to measure the degree of consistency of a logic theory. The purpose of the paper is to give a brief overview of the background, basic results and recent developments of quantitative logic.

Keywords: Quantitative logic, Σ-(α-tautology), Truth degree, Pseudo-metric, Entailment degree, Consistency degree.

1 Introduction

Classical two-valued logic is inadequate to face the essential vagueness of human reasoning which is approximate rather than precise in nature, although it fits quite well to mathematical reasoning. The main logical approach to treating with the concepts of vagueness and uncertainty is to establish systems of many-valued logic by accepting more truth values than the classical ones 1 and 0 only [1]. But logicians have not achieved an agreement on how to intuitively interpret these (additional) truth values. Rosser and Turquette [2] chose a subset D^+ of the set W of all truth values whose elements are called designated truth values to code degrees of truth of propositions, and another subset D^- of antidesignated truth values to code degrees of falsity. Of course one supposes that $1 \in D^+$, $0 \in D^-$ and $D^+ \cap D^- = \emptyset$. Having the designated and antidesignated truth values of a system of many-valued logic one then has an almost standard way to generalize the notions of tautology and of contradiction in classical logic. A formula A of the language of a system of propositional many-valued logic is a tautology iff its truth value always is a designated one under any valuation, and a formula B is a contradiction iff its truth value always is antidesignated. But, in modern systems

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 81–91.
springerlink.com

of many-valued logic, logicians usually choose 1 as the only designated truth value and 0 as the only antidesignated one [3-5]. It seems a little unreasonable that one uses only the classical truth values to code degrees of truth of propositions in a system of many-valued logic. Pavelka [6] abstracted propositions of Łukasiewicz $[0,1]$-valued logic to a fuzzy set with truth values as their membership degrees, and Novák *et al*[7] further extended this theory. However, this approach does not yet solve the problem which truth values code degrees of truth.

In order to consider additional truth values for generalizing the notion of tautology in classical logic, Wang [8] introduced the theory of Σ-(α-tautology) in a system of many-valued logic. More precisely, let α be a real number in the open-closed unit interval $(0,1]$, Σ a subset of the set Ω of all valuations, and A a proposition. Then A is called a Σ-(α-tautology) (*resp.* Σ-(α^+-tautology)) if $v(A) \geq \alpha$ (*resp.* $v(A) > \alpha$) for all $v \in \Sigma$, and a Σ-(α-tautology) A is said to be attainable if there exists $v \in \Sigma$ such that $v(A) = \alpha$. In particular, a Σ-(α-tautology) A is called a Σ-tautology if $\alpha = 1$, and A is called an α-tautology if $\Sigma = \Omega$. By duality, one can introduce the notion of Σ-(α-contradiction). Many interesting results have been obtained. For example, in the formal deductive system $\mathcal{L}^*$ (equivalently, in the nilpotent minimum logic), there are exactly three classes of α-tautologies, namely, tautologies (i.e., 1-tautologies), $(\frac{1}{2})^+$-tautologies and $\frac{1}{2}$-tautologies, and in Łukasiewicz $[0,1]$-valued propositional logic, there exists attainable r-tautology for every rational number $r \in (0,1]$. For more details one can also consult [9,10]. It is by evaluating the lower bound of truth values a proposition can take that this method generalized the notion of tautology. This seems too rough for classification of α-tautologies in $\mathcal{L}^*$, and even for an α-tautology A, it is fully possible that there is a valuation u such that $u(A) \geq \alpha + \varepsilon$ for some positive number ε small enough. This fact shows that α cannot actually measure the exact degree of truth of an α-tautology. Does there exist an effective way to measure the degrees of truth of formulas in systems of many-valued logic? It is interesting that an exact method for coding degrees of truth of formulas was first proposed in classical propositional logic by measuring the size of models of propositions [11]. We call the crucial notion truth degree in the present paper. Then the notion of truth degree of propositions in standard complete fuzzy propositional logics was proposed by the Lebesgue integral of truth functions [12]. Thus we established the theory of truth degrees of propositions in logic systems where the truth value sets jumped from $\{0,1\}$ to $[0,1]$. In order to harmoniously fill in the gap of theories of truth degree between classical and fuzzy logic systems, Wang *et al* [13] extended this notion to n-valued Łukasiewicz propositional logic such that the truth degree function in n-valued case converges into the one in fuzzy case as n turns into infinity. Hence, the harmonious theory of truth degree of propositions in many-valued propositional logics was successfully established. This idea has also been extended to modal logic [14] and to first-order predicate logic [15]. Based on the fundamental notion of truth degree, one can introduce, in a very natural way, the notions of similarity degree and of pseudo-metric between propositions to establish approximate reasoning theory, and one can also propose an effective way to measure degrees of consistency of theories. We call such a new branch of research Quantitative Logic. The aim of the paper is to provide a survey of Quantitative Logic, covering its basic results and recent developments.

2 Truth Degrees of Formulas

In this section we recall the notion of truth degree in proper order in classical propositional logic, in n-valued propositional logic, and in fuzzy propositional logic. We assume in this paper that the truth value sets of n-valued propositional logics are the sets of rational numbers between 0 and 1 of the form

$$W_n = \{0, \frac{1}{n-1}, \cdots, \frac{n-2}{n-1}, 1\}.$$

Even if we have defined the notion of truth degree in any standard complete many-valued propositional logic, we shall restrict ourselves to Łukasiewicz logic Ł, Gödel logic G, product logic Π and the formal deductive system $\mathcal{L}^*$, and their n-valued extensions $Ł_n$, G_n, Π_n and $\mathcal{L}^*_n$.

In order to better understand the definition of truth degree we should first recall a fundamental theorem concerning the infinite dimensional product of measure spaces [16].

Theorem 1. *Let $(X_k, \mathcal{A}_k, \mu_k)$ be a sequence of probability measure spaces. Then there exists a unique product probability measure μ on the σ-algebra $\mathcal{A}$ of $X = \prod_{k=1}^{\infty} X_k$, which is generated by all the sets of the form $A_1 \times \cdots \times A_m \times X_{m+1} \times X_{m+2} \times \cdots$ with $A_k \in \mathcal{A}_k, k = 1, 2, \cdots, m$ and $m \in N$, such that, for each measurable set of the kind $E \times X_{m+1} \times X_{m+2} \times \cdots$,*

$$\mu(E \times X_{m+1} \times X_{m+2} \times \cdots) = (\mu_1 \times \cdots \times \mu_m)(E). \tag{1}$$

In the case of $E = A_1 \times \cdots \times A_m$ with $A_k \in \mathcal{A}_k$ for any $k = 1, 2, \cdots, m$, one has

$$\mu(E \times \prod_{k=m+1}^{\infty} X_k) = (\mu_1 \times \cdots \times \mu_m)(E) = \mu_1(A_1) \times \mu_2(A_2) \times \cdots \times \mu_m(A_m). \tag{2}$$

2.1 Truth Degrees of Propositions in Classical Logic

Definition 1. *Let $X_k = W_2 = \{0, 1\}, \mathcal{A}_k = \mathcal{P}(X_k)$, and μ_k the uniform probability measure on X_k for all $k = 1, 2, \cdots$. Let $X = \prod_{k=1}^{\infty} X_k$ (i.e., $X = \Omega_2$ the set of all valuations), $\mathcal{A}$ the σ-algebra generated by $\mathcal{A}_1, \mathcal{A}_2, \cdots$ as in*

Theorem 1, and μ the unique product probability measure on $\mathcal{A}$ generated by $\mu_1, \mu_2, \cdots$. Define, for every proposition A in classical propositional logic,

$$\tau_2(A) = \mu(A^{-1}(1)), \tag{3}$$

where $A^{-1}(1)$ stands for the set $\{v \in \Omega_2 | v(A) = 1\}$ of models of A, then $\tau_2(A)$ is called the **truth degree** *of A.*

For each proposition $A = A(p_{i_1}, \cdots, p_{i_m})$, it is not difficult to check by (2) that $\tau_2(A) = \frac{1}{2^m} \mid \bar{A}^{-1}(1) \mid$, where $\bar{A}$ means the truth function (more precisely, Boolean function) represented by A and the symbol $|E|$ stands for the cardinal of E. This is just the expression given in the survey paper [17]. This expression is helpful to the calculation of (3). For example, $\tau_2(p) = \tau_2(q) = \frac{1}{2}, \tau_2(p \vee q) = \tau_2(p \to q) = \frac{3}{4}$ and $\tau_2(p_1 \vee \cdots \vee p_m) = 1 - \frac{1}{2^m}$.

Among interesting properties of $\tau_2(A)$ we select the following:

Proposition 1. *Let A, B be propositions. Then:*
(i) A is a tautology iff $\tau_2(A) = 1$,
(ii) A is a contradiction iff $\tau_2(A) = 0$,
(iii) $\tau_2(\neg A) = 1 - \tau_2(A)$,
(iv) $\tau_2(A) + \tau_2(B) = \tau_2(A \vee B) + \tau_2(A \wedge B)$,
(v) $\tau_2(B) \geq \tau_2(A) + \tau_2(A \to B) - 1$,
(vi) $\tau_2(A) + \tau_2(A \to B) = \tau_2(B) + \tau_2(B \to A)$.

Let

$$H_2 = \{\tau_2(A) \mid A \text{ is a proposition}\}.$$

Then we have:

Theorem 2. *([11]) $H_2 = \{\frac{k}{2^m} \mid k = 0, 1, \cdots, 2^m; m = 1, 2, \cdots\}$.*

2.2 Truth Degrees of Propositions in n-Valued Logics

Definition 2. *Let $X_k = W_n$, $\mathcal{A}_k = \mathcal{P}(X_k)$ and μ_k be the uniform probability measure on X_k for all $k = 1, 2, \cdots$, and let $X = \prod_{k=1}^{\infty} X_k = \Omega_n$ (the set of all valuations), $\mathcal{A}$ the σ-algebra generated by $\mathcal{A}_1, \mathcal{A}_2, \cdots$ and μ the unique product probability measure on $\mathcal{A}$ as in Theorem 1. Define, for every proposition A is a system of n-valued logic,*

$$\tau_n(A) = \sum_{i=0}^{n-1} \frac{i}{n-1} \mu(A^{-1}(\frac{i}{n-1})), \tag{4}$$

where $A^{-1}(\frac{i}{n-1})$ stands for the set of $\frac{i}{n-1}$-models of A, i.e., $A^{-1}(\frac{i}{n-1}) = \{v \in \Omega_n \mid v(A) = \frac{i}{n-1}\}$ for all $i = 0, 1, \cdots, n-1$. Then $\tau_n(A)$ is called the **truth degree** *of A.*

It is clear that $\tau_n(A) = \tau_2(A)$, defined by (3), for every A in the case of $n = 2$. Moreover, one can also check by (2) that $\tau_n(A) = \frac{1}{n^m} \sum_{i=0}^{n-1} \frac{i}{n-1} \mid \bar{A}^{-1}(\frac{i}{n-1}) \mid$ as given in [17] if A contains m atomic propositions, where $\bar{A}$ is the truth function induced by A.

It is easy to check that Proposition 1 still holds in the commonly used systems of n-valued propositional logic such as $Ł_n$, G_n, Π_n and $\mathcal{L}_n^*$ except the item (vi).

2.3 Truth Degrees of Propositions in Fuzzy Propositional Logic

In [0,1]-valued logics, the integral of truth functions can be employed to measure degrees of truth of the corresponding propositions [12].

Definition 3. *Let $A = A(p_{i_1}, \cdots, p_{i_m})$ be a proposition in a fuzzy logic, and $\bar{A} = \bar{A}(x_{i_1}, \cdots, x_{i_m})$ the induced truth function. Then the Lebesgue integral of $\bar{A}$ on $[0,1]^m$:*

$$\tau_\infty(A) = \int_{[0,1]} \bar{A}(x_{i_1}, \cdots, x_{i_m}) \mathrm{d}x_{i_1} \cdots \mathrm{d}x_{i_m} \tag{5}$$

is called the **(integrated) truth degree** *of A.*

For detailed properties of τ_∞ one can consult [12,3]. What we are interested in here is the relationship between the truth degree functions τ_n and τ_∞ when n turns into infinity. The following limit theorem obtained in [13] gives us a positive answer.

Theorem 3. *For every proposition A in Łukasiewicz logic, one has*

$$\lim_{n\to\infty} \tau_n(A) = \tau_\infty(A). \tag{6}$$

Indeed, the idea behind the proof of Theorem 3 can also be extended to Gödel logic, product logic and the logic $\mathcal{L}^*$, that is, the (6) holds also truth in these mentioned logics, see, e.g., [18].

2.4 Recent Developments

In Quantitative Logic, it is easy to check that $\tau_2(p) = \tau_n(p) = \tau_\infty(p) = \frac{1}{2}$ for all basic events p. This situation contradicts with the actual fact in our real life that the probabilities of different basic events may be greatly different. Keeping this in mind, Wang and Hui [19] introduced the notion of random truth degree of propositions in classical logic.

Definition 4. *Let $D = (P_1, P_2, \cdots)$ be a sequence of real numbers in (0,1), $x = (x_1, x_2, \cdots) \in \Omega_2$ a valuation. $\forall k = 1, 2, \cdots$, let $Q_k^{x_k} = P_k$ if $x_k = 1$ and $Q_k^{x_k} = 1 - P_k$ if $x_k = 0$, and define*

$$\tau_D(A) = \Sigma\{Q_1^{x_1} \times \cdots \times Q_n^{x_n} \mid (x_1, \cdots, x_n) \in \bar{A}^{-1}(1)\} \tag{7}$$

for every proposition $A = A(p_{i_1}, \cdots, p_{i_m})$*, where* $n = \max\{i_1, \cdots, i_m\}$*. Then* $\tau_D(A)$ *is called* D**-random truth degree** *of* A*.*

Only from the definition of (7) one cannot see the distinction and relationship between $\tau_D(A)$ and $\tau_2(A)$, but it actually randomizes the $\tau_2(A)$. In fact, for each coordinate P_k in D, P_k decides a probability measure μ_k on $X_k = \{0, 1\}$ with $\mu_k(\{1\}) = P_k = Q_k^1$ and $\mu_k(\{0\}) = 1 - P_k = Q_k^0$. Let μ be the unique product probability measure on $\Omega_2 = X = \prod_{k=1}^{\infty} X_k$ generated by μ_k's, and it is easy to check that

$$\tau_D(A) = \mu(A^{-1}(1)). \tag{8}$$

It follows from (8) and (3) that $\tau_D(A)$ and $\tau_2(A)$ are defined in the same way, which are both based on the product probability measure μ on the truth value set Ω_2 except that the μ in (8) is not necessarily uniform. In particular, if the probability measures $\mu_1, \mu_2, \cdots$ induced by the sequence D in (8) are all uniform, then it is obvious that $\tau_D(A) = \tau_2(A)$. But (8) still requires μ to be a product probability measure, and hence μ is independent, i.e., $\tau_D(p \wedge q) = \tau_D(p) \times \tau_D(q)$ for two different atomic propositions. This result is also unacceptable in particular applications. To overcome the difficulties the second author introduced in [20] the notion of probabilistic truth degree of propositions in classical logic by means of a Borel probability measure on Ω_2 (which is not necessarily uniform or independent), which can bring $\tau_2(A)$ and $\tau_D(A)$ as special cases into a unified framework. The idea can also be extended to systems of many-valued logic, see [21,22].

In the following, we assume that $X_k = W_2 = \{0, 1\}$ be the discrete topological space endowed with the usual topology, and hence that $\Omega_2 = X = \prod_{k=1}^{\infty} X_k = 2^\omega$ be the usual product topological space.

Definition 5. *Let* μ *be a Borel probability measure on the product topological space* $\Omega_2 = 2^\omega$*, and define*

$$\tau_\mu(A) = \mu(A^{-1}(1)) \tag{9}$$

for every proposition A *in classical logic. Then* $\tau_\mu(A)$ *is called the* μ-**probabilistic truth degree** *of* A*.*

Remark 1. (i) Without loss of generality, we assume now that every proposition we are dealing with is built up from the first m atomic propositions $p_1, \cdots, p_m$ for some m. Let $A = A(p_1, \cdots, p_m)$ be a proposition, and μ a Borel probability measure on Ω_2. If define $\mu(m) : \mathcal{P}(\{0, 1\}^m) \to [0, 1]$ by

$$\mu(m)(E) = \mu(E \times \prod_{k=m+1}^{\infty} X_k), \quad E \in \mathcal{P}(\{0, 1\}^m), \tag{10}$$

then $\mu(m)$ is a probability measure on the finite set $\mathcal{P}(\{0,1\}^m)$ and $\tau_\mu(A) = \mu(m)(\bar{A}^{-1}(1))$.

(ii) Let $A = A(p_1, \cdots, p_m)$ be a proposition and μ a product probability measure on Ω_2 generated by some probability measures μ_k on $X_k = \{0,1\}(k = 1, 2, \cdots)$. Then one has

$$\begin{aligned}\tau_\mu(A) &= \mu(m)(\bar{A}^{-1}(1)) = (\mu_1 \times \cdots \times \mu_m)(\bar{A}^{-1}(1)) \\ &= \Sigma\{\mu_1(\{x_1\}) \times \cdots \times \mu_m(\{x_m\}) \mid (x_1, \cdots, x_m) \in \bar{A}^{-1}(1)\} \\ &= \tau_D(A),\end{aligned}$$

where $D = (P_1, P_2, \cdots)$ and $P_k = \mu_k(\{1\})$ for all $k = 1, 2, \cdots$. Thus the random truth degree of formulas in (8) is only a special case of μ- probabilistic truth degree of formulas in (9) in the case where the μ is a product probability measure. In particular, if each μ_k is uniform, then $\tau_\mu(A) = \Sigma\{\frac{1}{2} \times \cdots \times \frac{1}{2} \mid (x_1, \cdots, x_m) \in \bar{A}^{-1}(1)\} = \frac{1}{2^m} \mid \bar{A}^{-1}(1) \mid = \tau_2(A)$.

(iii) Every valuation $v \in \Omega_2$ is a probabilistic truth degree function in the sense of Definition 5 where the involved Borel probability measure μ satisfies $\mu(E) = 1$ if $v \in E$ and otherwise $\mu(E) = 0$.

It is easy to check that Proposition 1 is true for finitely-atomic Borel probability measures μ, i.e., $\mu(m)(\{(x_1, \cdots, x_m)\}) \neq 0$ for every $(x_1, \cdots, x_m) \in \{0,1\}^m$ and for every $m \in N$. For every Borel probability measure μ on Ω_2, one can also define $H_\mu = \{\tau_\mu(A) \mid A$ is a proposition$\}$ and show that:

Theorem 4. *([22])(i) $H_\mu = \{\mu(m)(E) \mid E \subseteq \{0,1\}^m, m \in N\}$,*

(ii) If μ is non-atomic, then H_μ is dense in [0,1],

(iii) If μ is generated by uniform probability measures as in Definition 1, then $H_\mu = H_2 = \{\frac{k}{2^m} \mid k = 0, 1, \cdots, 2^m, m = 1, 2, \cdots\}$.

The truth degree functions defined by (3)–(5) and(7)–(9) can be used to introduce the graded versions of the notions of logical equivalence and of (semantic) entailment. In the following let us limit ourselves to the truth degree function τ_μ in (9).

3 Similarity Degrees between Propositions

Definition 6. *Let μ be a Borel probability measure on Ω_2, and A, B two propositions. Define*

$$\xi_\mu(A, B) = \tau_\mu((A \to B) \wedge (B \to A)), \tag{11}$$

*the $\xi_\mu(A, B)$ is called the μ-**similarity degree** between A and B.*

Clearly, the μ-similarity degree between logically equivalent propositions A and B is equal to 1, i.e., $\xi_\mu(A, B) = 1$ whenever A and B are logically equivalent. Conversely, if the μ is finitely-atomic, then one can check that $\xi_\mu(A, B) = 1$ is sufficient for logical equivalence between A and B. Then one has the following:

Proposition 2. *Let A, B and C be arbitrary propositions. Then:*
(i) $\xi_\mu(A, B) = \xi_\mu(B, A)$;
(ii) $\xi_\mu(A, B) + \xi_\mu(B, C) \leq \xi_\mu(A, C) + 1$;
(iii) $\xi_\mu(A, B) = 1 + \tau_\mu(A \wedge B) - \tau_\mu(A \vee B)$;
(iv) $\xi_\mu(A, B) + \xi_\mu(A, \neg B) = 1$;
(v) If μ is finitely-atomic, then

$$\xi_\mu(A, B) = 1 \text{ iff } A \text{ and } B \text{ are logically equivalent.}$$

(vi) If μ is finitely-atomic, then
$\xi_\mu(A, B) = 0$ *iff one of A and B is a tautology and the other is a contradiction.*

By virtue of the items (i) and (ii) of Proposition 2, one can introduce a pseudo-metric on the set $F(S)$ of all propositions.

Definition 7. *Let μ be a Borel probability measure on Ω_2, and A and B propositions. Define*

$$\rho_\mu(A, B) = 1 - \xi_\mu(A, B). \tag{12}$$

Then, by Proposition 2(i) and (ii), ρ_μ is a pseudo-metric on $F(S)$.

The pseudo-metric space $(F(S), \rho_\mu)$ has the following properties:

Theorem 5. *([22]) (i) If μ is non-atomic, then $(F(S), \rho_\mu)$ has no isolated points,*
(ii) For every Borel probability measure μ on Ω_2, logical connectives $\neg, \vee, \wedge$ and $\rightarrow$ are all uniformly continuous with respect to ρ_μ.

4 Entailment Degrees of Propositions from Theories

As mentioned above, the truth degree function (9) can be used to grade the notion of semantic entailment.

Let Γ be a theory and A a proposition in classical logic. Assume first that Γ semantically entails A, then, by the strong completeness theorem of classical logic, Γ syntactically entails A. It follows from the deduction theorem that there exists a finite sequence of propositions from Γ, say $A_1, \cdots, A_m$, such that $A_1 \wedge \cdots \wedge A_m \rightarrow A$ is a theorem. And consequently, one has $\tau_\mu(A_1 \wedge \cdots \wedge A_m \rightarrow A) = 1$. The inverse argument is also true provided that the μ is finitely atomic. From the above analysis, in order to determine whether a given proposition A is a semantic consequence of a theory Γ, it suffices to calculate the truth degrees $\tau_\mu(A_1 \wedge \cdots \wedge A_m \rightarrow A)$ of all possible propositions of the form $A_1 \wedge \cdots \wedge A_m \rightarrow A$ with $A_1, \cdots, A_m \in \Gamma$. Taking infinite theories into account, we use the supremum of truth degrees of these propositions to measure the degree of entailment of A from Γ. This idea is made more precisely by the following definition.

Definition 8. *Let Γ be a theory and A a proposition in classical logic. Let μ be a Borel probability measure on Ω_2. Then*

$$Entail_\mu(\Gamma, A) = \sup\{\tau_\mu(A_1 \wedge \cdots \wedge A_m \to A) \mid A_1, \cdots, A_m \in \Gamma, m \in N\} \quad (13)$$

*is called the μ-**entailment degree** of A from Γ.*

Example 1. Let μ be the product probability measure given in Definition 1. Find $\mathrm{Entail}_\mu(\Gamma, A)$, where
(i) $\Gamma = \{p\}, A = q$,
(ii) $\Gamma = \{p_2, p_3, \cdots\}, A = p_1$.

Solution 1. (i) By (12) and (13), $\mathrm{Entail}_\mu(\Gamma, A) = \tau_2(p \to q) = \frac{1}{2^2} \mid \overline{p \to q}^{-1}(1) \mid = \frac{3}{4}$.
(ii)

$$\begin{aligned} \mathrm{Entail}_\mu(\Gamma, A) &= \sup\{\tau_2(p_2 \wedge \cdots \wedge p_m \to p_1) \mid m \in N\} \\ &= \sup\{\tfrac{2^m - 1}{2^m} \mid m \in N\} \\ &= 1. \end{aligned}$$

This shows that A is a semantic consequence of Γ in the degree 1, but not a semantic consequence of Γ in its original sense.

Recall that a theory is said to be inconsistent if the contradiction $\bar{0}$ is its consequence. Hence $\mathrm{Entail}_\mu(\Gamma, \bar{0})$ is an ideal index to measure the degree of inconsistency of Γ. Perhaps this hints the idea that one may define the consistency degree of Γ to be 1-$\mathrm{Entail}_\mu(\Gamma, \bar{0})$, but this idea has a shortcoming that it could not distinguish consistent theories with $\mathrm{Entail}_\mu(\Gamma, \bar{0}) = 1$ from inconsistent ones as shown by Example 1(ii). So one has to slightly revise it.

Definition 9. *Let Γ be a theory in classical logic. Define*

$$Consist_\mu(\Gamma) = 1 - \frac{1}{2} Entail_\mu(\Gamma, \bar{0})(1 + i_\mu(\Gamma))$$

*where $i_\mu(\Gamma) = [\max\{1 - \tau_\mu(A_1 \wedge \cdots \wedge A_m) \mid A_1, \cdots, A_m \in \Gamma\}]$. Then $Consist_\mu(\Gamma)$ is called the μ-**consistency degree** of Γ.*

One can then prove the following:

Theorem 6. *Let Γ be a theory, and μ a finitely atomic Borel probability measure on Ω_2. Then:*
(i) Γ is inconsistent iff $Consist_\mu(\Gamma) = 0$,
(ii) Γ is consistent iff $\frac{1}{2} \leq Consist_\mu(\Gamma) \leq 1$,
(iii) Γ is consistent and $Entail_\mu(\Gamma) = 1$ iff $Consist_\mu(\Gamma) = \frac{1}{2}$.

For more results about consistency degrees of theories one can consult [23].

It is interesting that the function Entail_μ in (12) is closely related to the pseudo-metric function ρ_μ in (11).

Theorem 7. *(i)* $\mathrm{Entail}_\mu(\Gamma, \bar{0}) = \mathrm{div}_\mu(\Gamma)$, *where* $\mathrm{div}_\mu(\Gamma) = \sup\{\rho_\mu(A, B) \mid A, B \in D(\Gamma)\}$.

(ii) $\rho_\mu(A, D(\Gamma)) + \mathrm{Entail}_\mu(\Gamma, A) = 1$.

Proof. (i) $\mathrm{div}_\mu(\Gamma) = \sup\{\rho_\mu(A, B) \mid A, B \in D(\Gamma)\} = 1 - \inf\{\xi_\mu(A, B) \mid A, B \in D(\Gamma)\}$. Since $\xi_\mu(A, B) \geq \tau_\mu(A \wedge B) = \xi_\mu(T, A \wedge B)$ where T is a tautology, one then has $\mathrm{div}_\mu(\Gamma) = 1 - \inf\{\xi_\mu(T, A \wedge B) \mid A, B \in D(\Gamma)\} = 1 - \inf\{\tau_\mu(A \wedge B) \mid A, B \in D(\Gamma)\} = 1 - \inf\{\tau_\mu(A_1 \wedge \cdots \wedge A_m) \mid A_1, \cdots, A_m \in \Gamma\} = \mathrm{Entail}_\mu(\Gamma, \bar{0})$.

(ii)

$$\begin{aligned}\rho_\mu(A, D(\Gamma)) &= \inf\{\rho_\mu(A, B) \mid B \in D(\Gamma)\}\\ &\leq \inf\{\rho_\mu(A, A \vee B) \mid B \in D(\Gamma)\}\\ &= \inf\{1 - \tau_\mu((A \to A \vee B) \wedge (A \vee B \to A)) \mid B \in D(\Gamma)\}\\ &= \inf\{1 - \tau_\mu(B \to A) \mid B \in D(\Gamma)\}.\end{aligned}$$

On the other hand,

$$\begin{aligned}\rho_\mu(A, D(\Gamma)) &= \inf\{\rho_\mu(A, B) \mid B \in D(\Gamma)\}\\ &= 1 - \sup\{\xi_\mu(A, B) \mid B \in D(\Gamma)\}\\ &= 1 - \sup\{\tau_\mu((A \to B) \wedge (B \to A)) \mid B \in D(\Gamma)\}\\ &\geq 1 - \sup\{\tau_\mu(B \to A) \mid B \in D(\Gamma)\}.\end{aligned}$$

This shows that $\rho_\mu(A, D(\Gamma)) = 1 - \sup\{\tau_\mu(B \to A) \mid B \in D(\Gamma)\}$. It is obvious that $\sup\{\tau_\mu(B \to A) \mid B \in D(\Gamma)\} = \sup\{\tau_\mu(A_1 \wedge \cdots \wedge A_m \to A) \mid A_1, \cdots, A_m \in \Gamma\} = \mathrm{Entail}_\mu(\Gamma, A)$. Thus (ii) is true.

Acknowledgements. This work described here is supported by the National Natural Science Foundation of China (Grant No.10771129), the Natural Program for Basic Research of Shaanxi Province, China (Grant No. 2010JQ8020), the Fundamental Research Funds for the Central Universities (Grant No. GK200902048) and the Youth Foundation of Shaanxi Normal University (Grant No.200901010).

References

1. Gottwald, S.: A treatise on many-valued logics. Research Studies Press, Bldock (2001)
2. Rosser, J.B., Turquette, A.R.: Many-valued logics. North-Holland, Amsterdam (1952)
3. Wang, G.J., Zhou, H.J.: Introduction to mathematical logic and resolution principle. Alpha Science International Ltd., Oxford (2009)
4. Hájek, P.: Metamathematics of fuzzy logic. Kluwer, Dordrecht (1998)
5. Esteva, F., Godo, L.: Monoidal t-norm based logic: towards a logic for left-continuous t-norms. Fuzzy Sets and Systems 124, 271–288 (2001)
6. Pavelka, J.: On fuzzy logic I, II, III. Zeit Math. Logic Grundl. Math. 25, 45–52, 119–134, 447–464 (1979)

7. Novák, V., Perfilieva, I., Močkoř, J.: Mathematical principles of fuzzy logic. Kluwer, Dordrecht (1999)
8. Wang, G.J.: Theory of Σ-(α-tautologies) in revised Kleene system. Sci. China Ser. E 28, 146–152 (1998)
9. Wu, H.B.: Theory of generalized tautologies in the revised Kleene system. Sci. China Ser. E 32, 224–229 (2000)
10. Yang, X.B., Zhang, W.X.: Theory of generalized tautologies in Łukasiewicz logic system. J. Shaanxi Normal Univ. 26 6-9(Natural Science Ed.) (1998)
11. Wang, G.J., Fu, L., Song, J.S.: Theory of truth degrees of propositions in two-valued propositional logic. Sci. China Ser. A 31, 998–1008 (2001)
12. Wang, G.J., Leung, Y.: Integrated semantics and logic metric spaces. Fuzzy Sets and Systems 136, 71–91 (2003)
13. Wang, G.J., Li, B.J.: Theory of truth degrees of formulas in n-valued Łukasiewicz propositional logic. Sci. China Ser. E 35, 561–569 (2005)
14. Wang, G.J., Duan, Q.L.: Theory of (n) truth degrees of formulas in modal logic and a consistent theorem. Sci. China Ser. F 52, 70–83 (2009)
15. Wang, G.J., Qin, X.Y., Zhou, X.N.: An intrinsic fuzzy set on the universe of discourse of predicate formulas. Fuzzy Sets and Systems 157, 3145–3158 (2006)
16. Halmos, P.R.: Measure theory. Springer, New York (1955)
17. Wang, G.J., Zhou, H.J.: Quantitative logic. Inform. Sci. 179, 226–247 (2009)
18. Zhou, H.J., Wang, G.J., Zhou, W.: Consistency degrees of theories and methods of graded reasoning in n-valued R_0-logic (NM-logic). Int. J. Approx. Reasoning 43, 117–132 (2006)
19. Wang, G.J., Hui, X.J.: Randomization of classical inference patterns and its application. Sci. China Ser. F 50, 867–877 (2007)
20. Zhou, H.J.: Probabilistic truth degrees of formulas in classical propositional logic. In: Proceedings of the International Conference on Quantitative Logic and Quantification of Software. Global-Link Publisher, Hong-Kong (2009)
21. Zhou, H.J.: Borel probabilistic and quantitative logic (2011) (submitted for publication)
22. Zhou, H.J.: Probabilistic and quantitative logic. Ph.D Thesis, Shaanxi Normal Univ., Xi'an (2009)
23. Zhou, H.J., Wang, G.J.: Generalized consistency degrees of theories w.r.t. formulas in several standard complete logic systems. Fuzzy Sets and Systems 157, 2058–2073 (2006)

D-Completions of Net Convergence Structures

Weng Kin Ho, Dong-sheng Zhao, and Wen Shih Wee

Mathematics and Mathematics Education, National Institute of Education, Nanyang Technological University, Singapore
e-mail: wengkin.ho@nie.edu.sg, dongsheng.zhao@nie.edu.sg, wee0ws@gmail.com

Abstract. By extending Ershov's notion of a d-space from topological spaces to net convergence spaces, this paper details the d-completion of certain net convergence structures which are rich enough to support it. In particular, it is demonstrated that spaces which are embeddable into d-spaces which have iterated limits admit d-completions. The main result reported herein generalizes an existing procedure for d-completion of T_0 spaces.

Keywords: Net convergence, epitopological space, d-space, d-completion, dcpo.

1 Introduction

A directed subset D of a partially ordered set P is a non-empty set for which any finite subset $F \subseteq D$ has an upper bound in D. Being an obvious generalization of chains, directed sets capture succinctly the essence of convergence of sequences in general topology. Indeed, one can traced back to works like [9], where directed sets, together with nets, were used to determine topologies. Riding on its special affinity to convergence, directed sets were soon used to model phenomenon of approximations in computation. Consequently, partial orders which support the existence of suprema (in a way, viewed as limit points) of such sets emerge to become a salient notion in topology, domain theory and denotational semantics. These structures are called *directed complete posets*, abbreviated by *dcpo's*, have now become well-known in the community of theoretical computer science (particularly, programming semantics). One of the most important appearance of this concept is in the definition of the famous *Scott topology* invented by D.S. Scott in the late 1960's: "A subset U of a partially ordered set is open if it is upper and is inaccessible by *directed* suprema." (See [3,5].)

Valued as an indispensable condition of completeness in partial orders, it sparked off an active line of research for order theorists to provide 'canonical' *dcpo-completions* in the event that a poset fails to be directed-complete.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 93–110.
springerlink.com

For instance, the pioneering work of G. Markowsky [8] details a procedure of dcpo-completion via chains. Directedness and directed-completeness continued their pervasion from the realm of posets to that of T_0 spaces – the key link here being the specialization order of the underlying topology: $x \sqsubseteq_\tau y$ if and only if $x \in \mathrm{cl}(\{y\})$, where cl is the closure operator with respect to the topology τ. Along this topological line of development, O. Wyler [12] and Y.L. Ershov [2] had already studied extensively the d-spaces which are topological spaces behaving like dcpo's (i.e., directed complete with respect to the specialization order and in which the limit points of directed subsets are exactly their suprema). In particular, Ershov showed that every T_0 space X has its *d-completion* (i.e., a universal completion parallel to that of dcpo-completions for posets) – the smallest ambient d-space into which X can be embedded. The inverse limits construction employed by [8] and [2] is essentially a 'bottom-up' approach. Later, a 'top-down' approach was carried out in recent works of T. Fan and D. Zhao [14], and that of K. Keimel and J.D. Lawson [6], exploiting a new topology called D-topology.

In this paper, we proceed to investigate the existence of d-completion for a net convergence space, which is one of the structures that generalizes topological spaces. More precisely, we introduce a suitable notion of d-space in the context of net convergence spaces and show that there exists a certain class $\mathcal{C}$ of (net) convergence spaces whose structures are rich enough to admit (the corresponding) d-completions. It turns out that such a class $\mathcal{C}$ contains all T_0 topological spaces, thus generalizing the existing results of d-completions of T_0 topologies spaces and dcpo-completions of posets.

The subsequent structure of this paper is as follows. Section 2 sets the appropriate categorical stage to facilitate many kinds of completions that will take place in the present context of convergence spaces, and this is supported by a few well-known examples. Section 3 then ushers in the notion of convergence space and its related constructs. Niceness conditions form the subject of discourse in Section 4; and spaces in which iterated limits exist for Section 5. These concepts, as it turns out, will be exploited in the development of the theory of d-spaces in the context of net convergence structures in Section 6. This section then culminates with the main theorem which states some sufficiency condition for the existence of d-completions with respect to the definition of d-spaces which we had stated in preceding section.

2 Universal K-Fications in Topological Categories

Our main study centers around the several universal **K**-fications in the category of net convergence spaces. We choose to carry out this programme within the convenient framework of topological categories in the sense of [11]. The notion of a concrete category is an essential one with regards to this choice.

By a concrete category we mean a category **C** whose *objects* are structured sets, i.e., pairs (X, ξ) where X is a set and ξ is a **C**-structure on X, whose *morphisms* $f : (X, \xi) \longrightarrow (Y, \eta)$ are suitable maps between X and Y and whose composition law is the usual composition of maps – in other words: a category **C** together with a faithful (forgetful) functor $U : \mathbf{C} \to \mathbf{Set}$ where **Set** denotes the category of sets and maps.

Definition 2.1. *A concrete category* **C** *is called* topological *if it satisfies the following conditions:*

1. Existence of initial structures*:*
   For any set X and any family $(X_i, \xi_i)_{i\in I}$ of **C***-objects X_i's indexed by I and mappings $\xi_i : X \to X_i$ each indexed by I, there exists a unique* **C***-structure (X, ξ) which is initial with respect to $(X, f_i, (X_i, \xi_i), I)$, i.e., such that for any* **C***-object (Y, η) a map $g : (Y, \eta)\longrightarrow(X, \xi)$ is a* **C***-morphism if and only if for every $i \in I$ the composite map $f_i \circ g : (Y, \eta)\longrightarrow(X_i, \xi_i)$ is a* **C***-morphism.*
2. Fibre-smallness*:*
   For any set X, the **C***-fibre of X, i.e., the class of all* **C***-structures on X, is a set.*
3. Terminal separator property*:*
   For any set X with cardinality one, there exists precisely one **C***-structure on X.*

The concept of a full reflective subcategory allows one to deal with, in a coherent manner, different kinds of completions which we are about to embark on.

Definition 2.2. *A full subcategory* **K** *of a category* **C** *is called* reflective *if the inclusion functor* Incl *has a left adjoint R, which then is called a* reflector.

Equivalently, this is realized in the following way: For each **C**-object C, there exists a **K**-object $\tilde{C}$ and a **C**-morphism $r_C : C\longrightarrow\tilde{C}$ such that for each **K**-object D and each **C**-morphism $f : C\longrightarrow D$, there is a unique **K**-morphism $\overline{f} : \tilde{C}\longrightarrow D$ such that

$$\overline{f} \circ r_C = f.$$

We call the object $\tilde{C}$ the **K**-modification of C and the universal **C**-morphism $r_C : C\longrightarrow\tilde{C}$ the (universal) reflection. Several subcategories **K** of the topological categories we consider in this paper are crucially of the following two kinds:

Definition 2.3. *A* bireflective *subcategory* **K** *of a category* **C** *is a reflective subcategory in which for each* **C***-object C the reflection $r_C : C\longrightarrow\tilde{C}$ is a bimorphism, i.e., it is both a monomorphism and an epimorphism.*

Definition 2.4. *A full subcategory* **K** *of a category* **C** *is said to be* isomorphism-closed *if for any* **C***-object C and any* **K***-object K, whenever C is isomorphic to K then C is itself a* **K***-object.*

The reason for singling out these kinds of subcategories is that:

Theorem 2.5.(Theorem 2.2.12,[11])
Every bireflective and isomorphism-closed subcategory $\mathbf{K}$ *of a topological category* $\mathbf{C}$ *is topological.*

The above theorem comes handy in ensuring the existence of initial structures for certain subcategories of net convergence spaces we are considering in this paper.

3 Net Convergence Spaces

In this section, we introduce our main character – net convergence spaces. A pre-ordered set I (i.e., one with a reflexive and transitive relation $\sqsubseteq$) is said to be *directed* if for every pair $i_1,\ i_2 \in I$, there is always an $i_3 \in I$ such that $i_1,\ i_2 \sqsubseteq i_3$. A *net* in a set X is a mapping from a directed pre-order I to X. If we need to be more explicit about the elements of a net, we use the notation $(x_i)_{i \in I}$. Otherwise, we use Greek letters $\mu,\ \nu$. In particular, for any pre-order I and any fixed element $x \in X$, the *constant net* $(x) : I \to X,\ i \mapsto x$ is a net in X. For any directed subset D of a partially ordered set P (poset, for short), D defines a net $(x_d)_{d \in D}$ where $x_d := d$ for each $d \in D$. Given a set X, we use $\Psi(X)$ to denote the set of all nets in X.

Throughout our discussion, we make a more than casual use of the phrase 'eventually' in the following sense: If $P(x)$ is a property of the elements $x \in X$, we say that $P(x_j)$ holds *eventually* in a net $(x_j)_{j \in J}$ if there is a $j_0 \in J$ such that $\varphi(x_k)$ holds whenever $k \sqsupseteq j_0$.

We say that $(x_i)_{i \in I}$ is a subnet of $(y_j)_{j \in J}$, denoted by $(x_i)_{i \in I} \leq (y_j)_{j \in J}$, if

$$\forall j_0 \in J, \exists i_0 \in I, \forall i \sqsupseteq i_0, x_i \in T_{j_0}^{(y_j)}.$$

Here, $T_{j_0}^{(y_j)}$ denotes the set $\{y_j \mid j \sqsupseteq j_0\}$ and is called the j_0^{th}-tail of the net $(y_j)_{j \in J}$. Equivalently, $(x_i)_{i \in I}$ is a subnet of $(y_j)_{j \in J}$ if and only if for each $j \in J$, the net $(x_i)_{i \in I}$ is eventually in the j^{th}-tail of $(y_j)_{j \in J}$. Then $\leq$ is a pre-order on $\Psi(X)$ which will be called the subnet pre-order.

Remark 3.1

(1) In the literature of net convergence structures, there are different definitions of a subnet. In particular, Kelley in [7] defines $(x_i)_{i \in I}$ *to be a subnet of* $(y_j)_{j \in J}$ *if and only if there is a mapping* $h : I \to J$ *such that*

1. $x_i = y_{h(i)}$ *for all* $i \in I$*, and*
2. *for each* $j \in J$*, there is* $i_0 \in I$ *such that whenever* $i \sqsupseteq i_0$ *then* $h(i) \sqsupseteq j$*.*

Several authors, such as [1], consider additionally that h *is monotone. Our version, adopted from [10], is less general but yet natural enough to crucially allow the safe passage of certain arguments regarding subnets where other competing versions fail.*

(2) Note that we may have two distinct nets μ_1 *and* μ_2 *which are subnet of each other. We will call a net* μ *a proper subnet of net* ν *if* $\mu \leq \nu$ *and* $\nu \not\leq \mu$*.*

Definition 3.2. *A net is called an* ultranet *if it has no proper subnet.*

In other words, it is minimal with respect to the subnet pre-order. More precisely, a net ν is an ultranet if for any subnet ν' of ν, ν is also a subnet of ν'. Given a net μ in a set X, by the Hausdorff maximality principle, there is always an ultranet ν which is a subnet of μ. The abridged version of saying this is that ν is an '*ultra-subnet*' of μ.

Example 3.3. *Every constant net (x) is an ultranet since for every subnet $(y_j)_{j\in J}$ of (x), it holds that (x) is eventually in the i^{th}-tail of $(y_j)_{j\in I}$ as all of the y_j's are actually x.*

Proposition 3.4. *The following conditions are equivalent for a net $(x_i)_{i\in I}$ in a set X.*

(i) $(x_i)_{i\in I}$ *is an ultranet.*
(ii) *For any subset E of X, $(x_i)_{i\in I}$ is eventually in either E or $X\backslash E$.*
(iii) *For any subsets A and B, if $(x_i)_{i\in I}$ is eventually in $A\cup B$ then $(x_i)_{i\in I}$ is either eventually in A or in B.*

Proof. (i) $\implies$ (ii): Suppose not. Then for each $i\in I$, there exist $i_1,\ i_2 \sqsupseteq i\in I$ such that $x_{i_1}\in E$ and $x_{i_2}\in X\backslash E$. This gives rise to a *proper* subnet of $(x_i)_{i\in I}$ whose elements are respectively in E. However, this contradicts the assumption that E has no proper subnets.

(ii) $\implies$ (iii): Consider $X = A\cup B$ with no loss of generality and set $E = A$. Then $X\backslash E = B\backslash A$. Applying (ii), the net $(x_i)_{i\in I}$ will eventually be in E or in $X\backslash E$. Equivalently, the net is eventually in A or in $B\backslash A$. Thus the net is eventually in A or in B.

(iii) $\implies$ (i): Suppose that $(x_i)_{i\in I}$ has a proper subnet $(y_j)_{j\in J}$. We assume, without lose of generality, that it is not true that $(x_i)_{i\in}$ is in $A = \{y_j \mid j\in J\}$ eventually. Then $B = \{x_i \mid i\in I\}\backslash A \neq \emptyset$. Since $(x_i)_{i\in I}$ is in $A\cup B$ eventually and not in A eventually, by (iii), the net $(x_i)_{i\in I}$ is in B eventually. Assume that $T^{(x_i)}_{i_0}\subseteq B$ holds for some $i_0\in I$. Then $(y_j)_{j\in J}$ is in $T^{(x_i)}_{i_0}$, and hence in B eventually. But this is not possible because all y_j' are in A and A is disjoint from B. This contradiction shows that $(x_i)_{i\in I}$ must be an ultranet.

Having explained what nets are, we are ready for the following.

Definition 3.5. *By a* net convergence space*, we mean a pair $(X,\to)$ where X is a non-empty set and $\to$ a relation between the set $\Psi(X)$ and X, i.e., $\to\subseteq \Psi(X)\times X$ such that the following axioms are satisfied:*

1. *(CONSTANT NET)*
   For every $x\in X$, we always have $(x)\to x$.
2. *(SUBNET)*
   If $(y_j)_{j\in J}\leq (x_i)_{i\in I}$ and $(x_i)_{i\in I}\to x$, then $(y_j)_{j\in J}\to x$.

We write $(x_i)_{i\in I}\to x$ for $((x_i)_{i\in I}, x)\in\to$. Where there might be convergence structures derived from an existing one on a same set X, we always use $\to_X$ to denote the original one.

A function $f : (X, \rightarrow_X) \longrightarrow (Y, \rightarrow_Y)$ between net convergence spaces is said to be *continuous* if $(x_i)_{i \in I} \rightarrow_X x$ implies $(f(x_i))_{i \in I} \rightarrow_Y f(x)$. **NConv** denotes the category of net convergence spaces and continuous functions. A special kind of continuous map which is crucial in our theory of D-completion deserves special mention:

Definition 3.6. *A pre-embedding of a convergence space X to another Y is a continuous mapping $e : X \longrightarrow Y$ such that*

$$(x_i)_{i \in I} \rightarrow_X x \iff (e(x_i))_{i \in I} \rightarrow_Y e(x).$$

An embedding *is an injective pre-embedding.*

A convergence space $(X, \rightarrow)$ is said to be T_0 if for every $x,\ y \in X$, the following holds:

$$(x) \rightarrow y \wedge (y) \rightarrow x \iff x = y.$$

A few pathological examples must now be in place.

Example 3.7

1. *If (X, τ) is a topological space, then $(X, \rightarrow_\tau)$ is a convergence space, where $(x_i)_{i \in I} \rightarrow_\tau x$ if and only if the net $(x_i)_{i \in I}$ converges to x in the topological sense, i.e., for every open set U of (X, τ), $x \in U$ implies that $x_i \in U$ eventually. Such a convergence space is called a* topological *convergence space. Denote by Ω the* Sierpinski space, *i.e.,* $\{0, 1\}$. *Ω is the convergence space in which every net $(x_i)_{i \in I}$ converges to 0 and $(x_i)_{i \in I} \rightarrow 1$ if and only if $x_i = 1$ eventually. Then clearly $(\Omega, \rightarrow)$ is topological.*
2. *Let P be a poset. Define $(x_i)_{i \in I} \rightarrow_d x$ if and only if there is a directed subset $D \subseteq P$ such that $\bigsqcup D \sqsupseteq x$ and for each $d \in D$, $x_i \sqsupseteq d$ eventually. It is straightforward to verify that $(P, \rightarrow_d)$ is a convergence space. In [13], it is shown that P is topological if and only if P is a continuous poset.*
   In general, let $\mathcal{M}$ be a collection of subsets of poset P such that $\{x\} \in \mathcal{M}$ for each $x \in P$. Define $(x_i) \rightarrow_{\mathcal{M}} x$ if and only if there is $A \in \mathcal{M}$ such that $\bigsqcup A \sqsupseteq x$ and $x_i \sqsupseteq a$ eventually for each $a \in A$. Then $(X, \rightarrow_{\mathcal{M}})$ is a convergence space.
3. *Let (X, τ) be a topological space and $\sqsubseteq_\tau$ be the specialization order on X (i.e., $x \sqsubseteq_\tau y$ iff $x \in \mathrm{cl}(\{y\})$). Define $(x_i)_{i \in I} \rightarrow_c x$ if and only if $x \in \mathrm{cl}(\bigcup_{k \in I} T^{(x_i)}{}_k^{\downarrow})$, where $A^\downarrow := \{y \in X \mid \forall a \in A. y \sqsubseteq a\}$. Note that $A^\downarrow$ is in fact the set of all lower bounds of A in X. Then $(X, \rightarrow_c)$ is a convergence space, first defined in [4] for dcpos with their Scott topology.*
4. *A pre-metric space is a pair (X, ρ), where X is a non-empty set and $\rho : X \times X \longrightarrow [0, \infty)$ is a function satisfying $\rho(x, x) = 0$ for all $x \in X$. Now define $(x_i)_{i \in I} \rightarrow x$ if $(\rho(x_i, x))_{i \in I}$ converges to 0. Then $(X, \rightarrow)$ is a convergence space.*
5. *On $\mathbb{R}^2$ define $(x_i, y_i)_{i \in I} \rightarrow_1 (x, y)$ iff $(x_i)_{i \in I}$ converges to x in $\mathbb{R}$. Then $(\mathbb{R}^2, \rightarrow_1)$ is a convergence space. In general, if $f : X \longrightarrow Y$ is a mapping from a set X into a topological space Y, we define $(x_i)_{i \in I} \rightarrow_f x$, for*

net $(x_i)_{i\in I}$ and element x in X, if $(f(x_i))_{i\in I}$ converges to $f(x)$. Then $(X, \rightarrow_f)$ is a convergence space.

6. *For any poset P and a net $(x_i)_{i\in I}$ in P, define $(x_i)_{i\in I} \rightarrow_D x$ if there is a directed subset $E \subseteq P$ such that $\bigsqcup E = x$ and for each $e \in E, x_i \in E \cap \uparrow e$ eventually. Then $(P, \rightarrow_D)$ is a convergence space.*

Readers who are familiar with the theory of filter convergence would have realized the glaring similarities between the definitions of filter convergence spaces and net convergence spaces. In fact, one can show that there is a categorical adjunction between them. A given net $(x_i)_{i\in I}$ in X canonically induces a filter in X, denoted by $[(x_i)_{i\in I}]$, defined by

$$[(x_i)_{i\in I}] = \{A \subseteq X \mid \exists i_0 \in I. A \supseteq T_{i_0}^{(x_i)}\}.$$

In other words, $[(x_i)_{i\in I}]$ is the filter generated by the filter base consisting of the tails of $(x_i)_{i\in I}$.

Remark 3.8. With this notation, it is easy to see that $(x_i)_{i\in I} \leq (y_j)_{j\in J}$ if and only if $[(y_j)_{j\in J}] \subseteq [(x_i)_{i\in I}]$.

Conversely, a given filter $\mathcal{F}$ on X induces a canonical net $\hat{\mathcal{F}}$ defined as follows. Define $I_{\mathcal{F}} := \{(a, A) \mid (a \in A) \wedge (A \in \mathcal{F})\}$ and impose the following pre-order:

$$(a, F) \leq (b, G) \iff G \subseteq F.$$

The canonical net in X induced by the filter $\mathcal{F}$ is given by the mapping:

$$\hat{\mathcal{F}} : I_{\mathcal{F}} \longrightarrow X,\ (a, A) \mapsto a.$$

Denote the set of all filters on X by $\Phi(X)$ and order it by the sub-filter order $\leq$ (in fact, it is the reverse inclusion). Experts in Galois connections would immediately recognize the adjunction of pre-orders between $\Phi(X)$ and $\Psi(X)$, i.e.,

$$\alpha_X : \Phi(X) \to \Psi(X),\ \mathcal{F} \to \hat{\mathcal{F}}$$

is left adjoint to

$$\beta_X : \Psi(X) \to \Phi(X),\ (x_i)_{i\in I} \to [(x_i)_{i\in I}].$$

Such a pair of adjoint maps, natural in X, is in fact an e-p (embedding-projection) pair, i.e., $\alpha_X \circ \beta_X \leq \mathrm{id}_{\Psi(X)}$ & $\beta_X \circ \alpha_X = \mathrm{id}_{\Phi(X)}$. It can be shown that such e-p pairs induce a categorical adjunction between the category of filter convergence spaces **FConv** and that of net convergence spaces **NConv**, i.e., **FConv** $\dashv$ **NConv**. At the moment of writing, it is not known to the authors whether these two categories are equivalent.

Given any function $f : X \to Y$, one can define $f^+ : \Psi(X) \to \Psi(Y)$ by

$$f^+(x_i)_{i\in I} := (f(x_i))_{i\in I}$$

and also $f^- : \Psi(Y) \to \Psi(X)$ as follows:

$$f^{-}(y_j)_{j\in J} := \alpha_X \circ [f^{-1}(T_j^{(y_j)}) \mid j \in J].$$

Here, $[f^{-1}(T_j^{(y_j)}) \mid j \in J]$ denotes the filter generated by the filter base $\{f^{-1}(T_j^{(y_j)}) \mid j \in J\}$. For any $(x_i)_{i\in I} \in \Psi(X)$ and any $(y_j)_{j\in J} \in \Psi(Y)$, $f^{+}((x_i)_{i\in I}) \leq (y_j)_{j\in J} \iff (x_i)_{i\in I} \leq f^{-}((y_j)_{j\in J})$. This can be justified as follows:

$$\begin{aligned}
& f^{+}((x_i)_{i\in I}) \leq (y_j)_{j\in J} \\
\iff & (f(x_i))_{i\in I} \leq (y_j)_{j\in J} \\
\iff & \forall j \in J. \exists i_0 \in I. T_{i_0}^{(f(x_i))} \subseteq T_j^{(y_j)} \\
\iff & \forall j \in J. \exists i_0 \in I. f(T_{i_0}^{(x_i)}) \subseteq T_j^{(y_j)} \\
\iff & \forall j \in J. \exists i_0 \in I. T_{i_0}^{(x_i)} \subseteq f^{-1}(T_j^{(y_j)}) \\
\iff & (x_i)_{i\in I} \leq f^{-}((y_j)_{j\in J}).
\end{aligned}$$

NConv admits both initial and terminal structures for any given family of mappings. Hence products and coproducts of any collection of spaces exist. All of these properties are not surprising, thanks to the fact that **NConv** is a topological category.

The existence of initial structures guarantees the existence of subspaces. More precisely, a subspace X_0 of a space X is a subset of X equipped with the initial structure with respect to the set inclusion $\iota : X_0 \hookrightarrow X$. In particular, that $(X_0, \rightarrow_{X_0})$ is a subspace of $(X, \rightarrow_X)$ is equivalent to having:

$$(x_i)_{i\in I} \rightarrow_{X_0} x \iff (x_i)_{i\in I} \rightarrow_X x.$$

Of course, every subspace of a topological convergence space is topological.

Crucially, the category **NConv** also admits exponentials. Given two spaces X and Y, let $[X \to Y]$ be the set of all continuous mappings from X to Y. For a net $(f_i)_{i\in I}$ in $[X \to Y]$ and a $f \in [X \to Y]$, define $(f_i)_{i\in I} \to f$ iff for any $(x_k)_{k\in K} \rightarrow_X x$, one has

$$(f_i)_{i\in I} \cdot (x_k)_{k\in K} \to f(x),$$

where $(f_i)_{i\in I} \cdot (x_k)_{k\in K}$ is defined to be the net $(f_i(x_k))_{(i,k)\in I\times K}$. It turns out that $([X \to Y], \to)$ is a net convergence space, which we call the *function space* from X to Y. Further to the existence of products and exponentials, one of course expects nothing less than the fact that **NConv** is a cartesian closed category. This is one of the major reasons why convergence spaces are considered in preference to topological spaces since it is well known that the category of topological spaces is not Cartesian closed.

Regarding function spaces, the following property is frequently used.

Proposition 3.9. *Let X and Y be convergence spaces. Then Y is homeomorphic to a subspace of $[X \to Y]$.*

Proof. Define $k : Y \longrightarrow [X \to Y]$ by $k(y) = \overline{y}$, the constant mapping $x \to y$. Clearly, each $k(y)$ is in $[X \to Y]$ and k is injective. To show that k is a continuous map from Y to $[X \to Y]$, suppose $(y_j)_{j\in J} \to_Y y$. We aim to show that $(k(y_j))_{j\in J} \to_{[X\to Y]} k(y)$. To achieve this, take any $(x_i)_{i\in I} \to_X x$. Now for each $(j,i) \in J \times I$, the term $(k(y_j))(x_i) = \overline{y_j}(x_i) = y_j$. But $(y_j)_{j\in J} \to_Y y$ by our supposition. So $(k(y_j)(x_i))_{(j,i)\in J\times I} \to_Y y = \overline{y}(x)$. This is, of course, equivalent to saying $(k(y_j))_{j\in J} \to_{[X\to Y]} k(y)$.

To show that k is an embedding, suppose the net $(y_j)_{j\in J}$ in Y is such that $(k(y_j))_{j\in J} \to_{[X\to Y]} k(y)$. Now apply the net $(k(y_j))_{j\in J}$ to the constant net (x) which, we know, always converges to x in X. Thus by the definition of the convergence structure on $[X \to Y]$, it follows that

$$(k(y_j))_{j\in J}(x) = (\overline{y_j}(x))_{j\in J} = (y_j)_{j\in J} \to_Y k(y)(x) = y.$$

We conclude that $(y_j)_{j\in J} \to_Y y$ which implies that k is an embedding.

In what follows, by a *(convergence) space* we shall always mean a net convergence space unless otherwise stated. Given a space $(X, \to)$, there is a *induced topology* on X, making X to be a topological space denoted by $TX := (X, \tau_{\to})$ and is defined as follows: A subset $U \in \tau_{\to}$ if for any net $(x_i)_{i\in I}$ in X, $(x_i)_{i\in I} \to x$ and $x \in U$ always imply that $x_i \in U$ eventually.

It is straightforward to check that τ is indeed a topology on X. Moreover, if $(x_i)_{i\in I} \to x$, then $(x_i)_{i\in I}$ converges to x topologically. For any poset P, the induced topology on $(P, \to_d)$ is the Scott topology($U \subseteq P$ is Scott open iff $U = \uparrow U$ and if $D \subseteq P$ is a directed set with $\bigsqcup D \in U$ then $D \cap U \neq \emptyset$).

Such an induced topology can be viewed in terms of categorical adjunctions. Denoting the category of topological spaces by **Top**, one has the following sequence of adjunctions:

$$\mathbf{Top} \dashv \mathbf{FConv} \dashv \mathbf{NConv}.$$

In this perspective, one can see T as the right adjunction of the composition of adjunctions. Consequently, one views net convergence space as a generalization of filter convergence space inasmuch as filter convergence space is that of topological space.

Given a topological space (X, τ), the specialization order on X is the partial order $\sqsubseteq_\tau$ (or $\sqsubseteq$ for short) defined by $x \sqsubseteq_\tau y$ iff $x \in \mathrm{cl}(\{y\})$. For a convergence space $(X, \to)$, we call the specialization order of $(X, \tau_{\to})$ the *specialization order* of $(X, \to)$.

Remark 3.10

1. *Every continuous map $f : (X, \to_X) \longrightarrow (Y, \to_Y)$ is also continuous with respect to the induced topology and hence monotone with respect to the specialization order.*
2. *If Y is a topological space and X is a subspace of Y, then the specialization order on X coincides with the restricted order of that on Y, that is, $x \sqsubseteq x'$ holds in X if and only if $x \sqsubseteq x'$ holds in Y.*

Just before we end this section, we want to make record of the following simple but useful result.

Proposition 3.11. *Every reflective subcategory of* **NConv** *whose reflector is the identity is isomorphism-closed.*

Proof. Let $\mathrm{id}_X : X \to \tilde{X}$ be the reflector on a convergence space $(X, \to_X)$. Given any **K**-object Y which is isomorphic to X via a homeomorphism $f : X \longrightarrow Y$, one can extend it to a unique continuous mapping whose underlying map is again $f : \tilde{X} \longrightarrow Y$. Thus for any convergent net $(x_i)_{i \in I} \to_{\tilde{X}} x$, it follows from the continuity of $f : \tilde{X} \longrightarrow Y$ that $(f(x_i))_{i \in I} \to_Y f(x)$. Passing this convergence through the continuous map $f^{-1} : Y \longrightarrow X$, one immediately has $(x_i)_{i \in I} \to_X x$. Since $\tilde{X}$ is coarser than X, it follows that $\to_X = \to_{\tilde{X}}$. Thus $(X, \to_X)$ is an object in **K**.

4 Niceness Conditions

Niceness is a salient aspect of net convergence spaces. Various niceness conditions, introduced in this section, correspond to existing one for filter convergence structures already considered by R. Heckmann in [4].

In the following, the order on a convergence space refers to the specialization order unless otherwise stated.

Definition 4.1. *A space* $(X, \to)$ *is* up-nice *if whenever* $(x_i)_{i \in I} \to x$ *and* $(y_j)_{j \in J}$ *is a net that satisfies the following condition (†):*

$$\forall i_0 \in I . \exists j_0 \in J . \forall j \sqsupseteq j_0 . y_j \in \uparrow T_{i_0}^{(x_i)} . \qquad (\dagger)$$

then $(y_j)_{j \in J} \to x$.

A space is down-nice *if* $(x_i)_{i \in I} \to x$ *and* $y \sqsubseteq x$ *then* $(x_i)_{i \in I} \to y$. *A space is* order-nice *if it is both up-nice and down-nice. The objects of order-nice spaces and morphisms of continuous maps together constitute a full subcategory of convergence space which is denoted by* **OnConv**.

Example 4.2

1. *Every topological convergence space is order-nice.*
2. *For any poset* P*, the space* $(P, \to_d)$ *is order-nice.*
3. *For a poset* P*, in general, the space* $(P, \to_D)$ *mentioned in Example 3.7(6) is neither up-nice nor down-nice.*

Theorem 4.3. *For a given space* $(X, \to)$*, define a coarser space* $N(X)$ *as follows:*

$$(x_i)_{i \in I} \to_{N(X)} x \iff \exists (z_k)_{k \in K} \in \Psi(X) . \forall k_0 \in K . \exists i_0 \in I . \forall i \sqsupseteq i_0 . x_i \in \uparrow T_{k_0}^{(z_k)} \qquad (\ddagger)$$

and $(z_k)_{k \in K} \to z$ *for some* $z \sqsupseteq x$*. Then* $N(X)$ *carries the finest order-nice convergence structure on the same set which is coarser than* X*.*

Proof. We now verify that $N(X)$ is an order-nice convergence space in stages:

1. From the definition of $\to_{N(X)}$, it's clear that if $(x_i)_{i\in I} \to x$ then $(x_i)_{i\in I} \to_{N(X)} x$. Thus the CONSTANT NET axiom is satisfied.
2. Assume that $(x_i)_{i\in I} \to_{N(X)} x$ and $(y_j)_{j\in J} \leq (x_i)_{i\in I}$. We want to show that $(y_j)_{j\in I} \to_{N(X)} x$. Since $(x_i)_{i\in I} \to_{N(X)} x$, one has a net $(z_k)_{k\in K}$ which satisfies the above condition (‡) at one's disposal. Now for each $k_0 \in K$, one can find $i_0 \in I$ such that whenever $i \sqsupseteq i_0$ it holds that $x_i \in \uparrow T_{k_0}^{(z_k)}$. Corresponding to this i_0, there is $j_0 \in J$ such that whenever $j \sqsupseteq j_0$ we have $y_j \in T_{i_0}^{(x_i)}$. Consequently, this means that if $j \sqsupseteq j_0$ then $y_j \in \uparrow T_{k_0}^{(z_k)}$. Thus $(y_j)_{j\in J} \to_{N(X)} x$.
3. We show now that $N(X)$ is an order-nice space. That $N(X)$ is down-nice follows trivially from the transitivity of $\sqsubseteq$. It suffices to show that it is up-nice. For that purpose, suppose that $(x_i)_{i\in I} \to_{N(X)} x$ and $(y_j)_{j\in J}$ is a net that satisfies (†). We aim to show that $(y_j)_{j\in J} \to_{N(X)} x$, i.e., there is a net $(z_k)_{k\in K}$ satisfying the property that for each $k_0 \in K$, there exists $j_0 \in J$ such that whenever $j \sqsupseteq j_0$ one has $y_j \in \uparrow T_{k_0}^{(z_k)}$. Since $(x_i)_{i\in I} \to_{N(X)} x$, there is already a net $(z_k)_{k\in K}$ such that
$$\forall k_0 \in K. \exists i_0 \in I. \forall i \sqsupseteq i_0. x_i \in \uparrow T_{k_0}^{(z_k)}$$
and $(z_k)_{k\in K} \to z$ with $z \sqsupseteq x$. Because of (†), for such an $i_0 \in I$, one can find $j_0 \in J$ such that whenever $j \sqsupseteq j_0$ then $y_j \sqsupseteq x_{i'}$ for some $i' \sqsupseteq i_0$. Now since $i' \sqsupseteq i_0$, it follows that $x_{i'} \sqsupseteq z_{k'}$ for some $k' \sqsupseteq k_0$. Consequently, this means that for each $k_0 \in K$, one can find $j_0 \in J$ such that whenever $j \sqsupseteq j_0$ then $y_j \sqsupseteq z_{k'}$ with $k' \sqsupseteq k_0$, i.e., $y_j \in \uparrow T_{k_0}^{(z_k)}$.
4. Finally, we show that $N(X)$ carries the finest convergence structure on the same set X which is coarser than X. Suppose that $(X, \to')$ is order-nice and coarser than $(X, \to)$. We want to show that $N(X)$ is finer than $(X, \to')$. To this end, assume that $(x_i)_{i\in I} \to_{N(X)} x$. By definition, there exists $(z_k)_{k\in K}$ such that for each $k_0 \in K$ there exists $i_0 \in I$ such that whenever $i \sqsupseteq i_0$ we have $x_i \in \uparrow T_{k_0}^{(z_k)}$ and and $(z_k)_{k\in K} \to z$ for some $z \sqsupseteq x$. Since $(X, \to')$ is coarser than $(X, \to)$ and $(X, \to')$ is down-nice, it follows that $(z_k)_{k\in K} \to' x$. Since $(X, \to')$ is up-nice, it follows that $(x_i)_{i\in I} \to' x$ by definition. Thus we can conclude that $N(X)$ is finer than $(X, \to')$ and the proof is complete.

Definition 4.4. *For a given space $(X, \to)$, we define $N(X)$ to be the* order-nice modification *of X.*

Proposition 4.5. **OnConv** *is a bireflective and isomorphism-closed subcategory of* **NConv**.

Proof. We extend N to a functor, assigning to each convergence space X its order-nice modification $N(X)$ and leaving each continuous map $f : X \to Y$ as it is. That $f : (X, \to_{N(X)}) \longrightarrow (Y, \to_{N(Y)})$ is continuous follows from the fact that f is monotone. Clearly, the identity mapping $\mathrm{id}_X : X \to N(X)$ is continuous. We aim to show that for each order-nice space $(Y, \to_Y)$ and continuous mapping $f : (X, \to_X) \longrightarrow (Y, \to_Y)$, the same mapping f is also a continuous mapping from $(X, \to_{N(X)})$ to $(Y, \to_Y)$. To this end, let $(x_i)_{i \in I} \to_{N(X)} x$. By definition, there is a net $(z_k)_{k \in K} \to_X z$ for some $z \sqsupseteq x$ such that

$$\forall k_0 \in K. \exists i_0 \in I. \forall i \sqsupseteq i_0. x_i \in \uparrow T_{k_0}^{(z_k)}.$$

Since $f : X \longrightarrow Y$ is continuous, we have $(f(z_k))_{k \in K} \to_Y f(z)$. Also f being continuous is monotone with respect to the specialization order so that $f(z) \sqsupseteq f(x)$. Moreover, for each $k_0 \in K$, there is $i_0 \in I$ such that whenever $i \sqsupseteq i_0$ we have $f(x_i) \in \uparrow T_{k_0}^{(f(z_k))}$. But Y is up-nice so that one can now conclude that $(f(x_i))_{i \in I} \to_Y f(z)$. Since $f(z) \sqsupseteq f(x)$ and Y is down-nice, we have that $(f(x_i))_{i \in I} \to_Y f(x)$ as desired.

Lastly, since the reflector is the identity, it follows from Proposition 3.11 that **OnConv** is isomorphism-closed.

The above proposition, together with Theorem 2.5, establishes the following fact:

Corollary 4.6. *Let $\{X_i \mid i \in I\}$ be a family of order-nice (resp. up-nice, down-nice) spaces and X a convergence space which carries the initial convergence structure with respect to a collection of functions $\{f_i : X \to X_i \mid i \in I\}$. Then X is order-nice (resp. up-nice, down-nice). In particular, all subspaces and products of order-nice (resp. up-nice, down-nice) spaces are themselves order-nice (resp. up-nice, down-nice).*

Proposition 4.7. *Let X and Y be convergence spaces. Then the following are equivalent:*

(i) Y is order-nice (resp. up-nice, down-nice).
(ii) $[X \to Y]$ is order-nice (resp. up-nice, down-nice).

Proof. We prove the above only for up-niceness. Down-niceness is similar.
(i) $\implies$ (ii): Suppose that $(f_i)_{i \in I} \to_{[X \to Y]} f$ and there is a net $(g_j)_{j \in J}$ such that

$$\forall i_0 \in I. \exists j_0 \in J. \forall j \sqsupseteq j_0. g_j \in \uparrow T_{i_0}^{(f_i)}.$$

We want to show that $(g_j)_{j \in J} \to_{[X \to Y]} f$. To achieve this, take an arbitrary $(x_k)_{k \in K} \to_X x$. Since $(f_i)_{i \in I} \to_{[X \to Y]} f$, it follows that $(f_i(x_k))_{(i,k) \in I \times K} \to_Y f(x)$. By the continuity of f_i and g_j, and their monotonicity, it follows that

$$\forall k_0 \in K. \forall i_0 \in I. \exists j_0 \in J. \forall j \sqsupseteq j_0. g_j(x_k) \in \uparrow T_{(i_0, k_0)}^{(f_i(x_k))}.$$

So this means that

$$\forall (i_0, k_0) \in I \times K. \exists (j_0, k_0) \in J \times K. \forall (j, k) \sqsupseteq (j_0, k_0). g_j(x_k) \in \uparrow T_{(i_0, k_0)}^{(f_i(x_k))}.$$

Now we appeal to the up-niceness of Y to conclude that $(g_j(x_k))_{(j,k)\in J\times K} \to_Y f(x)$. Thus $(g_j)_{j\in J} \to_{[X\to Y]} f$.
(ii) $\Longrightarrow$ (i): Since Y may be regarded as a subspace of the order-nice space $[X \to Y]$, it immediately follows from Proposition 3.9 that Y is up-nice.

Corollary 4.8. *For any space X, the space $\Omega X := [X \to \Omega]$ is always order-nice (resp. up-nice, down-nice).*

Theorem 4.9. **OnConv** *is a Cartesian closed category.*

Proof. Because (1) the subcategory **OnConv** is closed under products and function spaces formed in **NConv**, and (2) the mappings are the same under the reflector N, the result then follows immediately from the Cartesian closedness of **NConv**.

5 Iterated-Limit Spaces

Kelley [7] characterized all topological convergence spaces to be exactly those convergence spaces which satisfy the following axioms:

1. (ITERATED LIMIT) A space $(X, \to)$ satisfies the ITERATED LIMIT (IL for short) axiom, if $(x_i)_{i\in I} \to x$ and $(x_{i,j})_{j\in J(i)} \to x_i$ for each $i \in I$, then $(x_{i,f(i)})_{(i,f)\in I\times M} \to x$, where $M = \Pi_{i\in I} J(i)$ is a product of directed sets.
2. (DIVERGENCE) If $(x_i)_{i\in I} \not\to x$, then $(x_i)_{i\in I}$ has a subnet $(y_j)_{j\in J}$ no subnet $(z_k)_{k\in K}$ of which ever has $(z_k)_{k\in K} \to x$.

In this section, we study the properties of spaces which satisfy the ITERATED LIMIT (IL, for short) axiom. We call such spaces *IL spaces.*

Remark 5.1. *It follows immediately that every topological space is an IL space.*

Proposition 5.2

1. *Let $\{X_i \mid i \in I\}$ be a family of IL spaces and X a convergence space which carries the initial convergence structure with respect to a collection of functions $\{f_i : X \to X_i \mid i \in I\}$. Then X is an IL space. In particular, all subspaces and products of IL spaces are themselves IL spaces.*
2. *Every retract of an IL space is an IL space.*

Proof. For (1), let $(x_k)_{k\in K} \to_X x$ with respect to the initial convergence and suppose for each $k \in K$, there is a directed set $J(k)$ such that $(x_{k,j})_{j\in J(k)} \to_X x_k$. We want to show that $(x_{k,g})_{(k,g)\in K\times M} \to_X x$ where $M = \prod_{k\in K} J(k)$. This means that we have to show that for each $i \in I$, it holds that

$$(f_i(x_{k,g(k)}))_{(k,g)\in K\times M} \to_{X_i} f_i(x).$$

Because for each $k \in K$ it holds that $(x_{k,j})_{j\in J(k)} \to_X x_k$, it holds that for each $k \in K$ and for each $i \in I$, one has $(f_i(x_{k,j}))_{j\in J(k)} \to_{X_i} f_i(x_k)$. Since X_i's are all IL spaces, this implies that for each $i \in I$, the following convergence holds:

$$(f_i(x_{k,g(k)}))_{(k,g)\in K\times M} \to_{X_i} f_i(x)$$

where $M = \prod_{k\in K} J(k)$, as desired.

For (2), assume that that $f : Y \longrightarrow X$, $g : X \to Y$ are continuous maps such that $f \circ g = \mathrm{id}_X$ and Y is an IL space. We proceed to show that X is also an IL space. For that purpose, assume that $(x_i)_{i\in I} \to_X x$ and for each $i \in I$, there is a directed set $J(i)$ such that $(x_{j,i})_{j\in J(i)} \to_X x_i$. Since g is continuous, one has for each $i \in I$ that $(g(x_{j,i}))_{j\in J(i)} \to_Y g(x_i)$. As Y is an IL space, it follows that $(g(x_{i,f(i)}))_{(i,f)\in I\times M} \to_Y g(x)$ where $M = \prod_{i\in I} J(i)$. Now the proof will be complete by applying the continuous f to this net and then invoking the fact that $f \circ g = \mathrm{id}_X$. Thus X is an IL space.

6 *d*-Spaces

For any net $(x_i)_{i\in I}$ in a space $(X, \to)$, denote

$$\lim(x_i)_{i\in I} = \{x \in X : (x_i)_{i\in I} \to x\}.$$

Definition 6.1. *A convergence space $(X, \to)$ is called a d-space if*

1. *it is order-nice and T_0,*
2. *for any directed subset $(x_i)_{i\in I}$ of X, $x = \bigsqcup_{i\in I} x_i$ exists and $(x_i)_{i\in I} \to x$, and*
3. *for every net $(x_i)_{i\in I}$, the limit set $\lim(x_i)_{i\in I}$ is closed under directed suprema.*

Example 6.2

1. *A topological space is a d-space if and only if it is a monotone convergence space.*
2. *For each poset P, $(P, \to_D)$ is a d-space if and only if $(P, \to_d)$ is a d-space, and in turn, if and only if P is a dcpo.*

Proposition 6.3

1. *If X is a d-space, then every closed set F of X is closed under taking supremum of directed sets.*
2. *If X is a d-space and $f : X \longrightarrow Y$ is a continuous mapping, then for any directed set D of X,*
$$f(\bigsqcup D) = \bigsqcup f(D).$$
3. *d-space is stable under retract.*

4. *Let X and Y be convergence spaces. Then $[X \to Y]$ is a d-space if and only if Y is a d-space. In particular, ΩX is a d-space for any convergence space X.*
5. *The product of any collection of d-spaces is again a d-space.*

Proof. 1. This follows from that every directed subset, as a net, converges to its supremum.

2. For any directed set $D \subseteq X$, $x = \bigsqcup D$ exists. Since f is monotone, it follows that $f(\bigsqcup D)$ is an upper bound of $f(D)$. If $y \in Y$ is an upper bound of $f(D)$, then $D \subseteq f^{-1}(\downarrow y)$. However, $\downarrow y = cl(\{y\})$ is a closed set, thus by part one, $\bigsqcup D \in f^{-1}(\downarrow y)$ and thus $f(\bigsqcup D) \leq y$. Hence $f(\bigsqcup D) = \bigsqcup f(D)$.
3. Let $s : X \longrightarrow Y$ and $r : Y \longrightarrow X$ be continuous mappings, where $r \circ s = \mathrm{id}_X$, X is a convergence space and Y a d-space. Since order-niceness has already been shown to be stable under retracts and trivially X is T_0, it remains to show that the convergence criteria are satisfied. Given any directed net $(x_i)_{i\in I}$ we wish to show that the supremum of $\{x_i : i \in I\}$ exists and $(x_i)_{i\in I} \to_X x$. Owing to the fact that Y is a d-space, it follows that the directed net $s(x_i)_{i\in I} \to_Y y$, where $y = \bigsqcup s(x_i)_{i\in I}$. Then by 2, $r(y) = \bigsqcup\{r(s(x_i)) : i \in I\} = \{x_i : i \in I\}$. Also $(x_i)_{i\in I} = (r(s(x_i)))_{i\in I} \to_X r(y)$. That $\lim(x_i)_{i\in I}$ is closed under directed sups follows from the continuity of r and s and the fact that for any directed set D of X, $\bigsqcup D = r(\bigsqcup s(D))$ proved just now.
4. It's easily shown that $[X \to Y]$ is T_0. By virtue of (3) and Proposition 3.9, it suffices to prove the sufficiency condition. So assume that Y is a d-space. Let $(f_i)_{i\in I}$ be a directed family in $[X \to Y]$. We propose that the function

$$f : X \longrightarrow Y, f(x) = \bigsqcup_{i\in I} f_i(x)$$

is the supremum of $(f_i)_{i\in I}$. Firstly, f is well-defined since $(f_i)_{i\in I}$ is directed with respect to the pointwise order and Y is a d-space so that the directed set $\{f_i(x) \mid i \in I\}$ in Y has a supremum. Next, we show that f is continuous. For any net $(x_k)_{k\in K}$ in X which converges to x, we aim to prove that $(f(x_k))_{k\in K} \to_Y f(x)$. Notice that $f(x_k) \sqsupseteq f_i(x_k)$ for each $i \in I$ so that by the order-niceness of Y and the continuity of the f_i's, one has $(f(x_k))_{k\in K} \to_Y f_i(x)$. Since $\lim(f(x_k))_{k\in K}$ is closed under directed suprema, it holds that $(f(x_k))_{k\in K} \to_Y \bigsqcup_{k\in K} f_k(x)$, i.e., $(f(x_k))_{k\in K} \to_Y f(x)$. Thus f is continuous. That f is indeed the supremum of the continuous mappings f_i's is obvious. Following the definition of the convergence in $[X \to Y]$ one can also verify straight forwardly that for any net $(h_i)_{i\in I}$ in $[X \to Y]$, $\lim(h_i)_{i\in I}$ is closed under directed suprema. Thus $[X \to Y]$ is a d-space.
5. We just show the second condition of d-space is satisfied by product of d-spaces. Let $(m_j)_{j\in J}$ be a directed family in $\prod_{i\in I} X_i$ where X_i's are d-spaces. Then for each $i \in I$, the image of the i-th projection $(\pi_i(m_j))_{j\in J}$

is a directed family in X_i since π_i is continuous. So $\bigsqcup_{j\in J}\pi_i(m_j)$ exists as X_i is a d-space. By putting $x_i := \bigsqcup_{j\in J}\pi_i(m_j)$ and define the net $x = (x_i)_{i\in I}$, then one clearly has $\bigsqcup_{j\in J} m_j = x$ and $(m_j)_{j\in J} \to x$.

Theorem 6.4. *The category of d-spaces is Cartesian closed.*

A subspace X_0 of a space X is called a d-base of X if for any $x \in X$, there is a directed family $\{y_i\}_{i\in I} \subseteq X_0$, such that $(y_i)_{i\in I} \to_X x = \bigsqcup\{y_i : i \in I\}$.

Central to the theory of D-completion of convergence spaces is the following fundamental lemma.

Lemma 6.5. *Let X be an order-nice space satisfying the IL axiom and X_0 be a d-base of X. Then every continuous mapping f_0 from the subspace X_0 into a d-space Y has a unique continuous extension f on X.*

Proof. For each $x \in X$, let $(y_i)_{i\in I}$ be a directed family in X_0 such that $(y_i)_{i\in I} \to_X \bigsqcup_{i\in I} y_i = x$. Define $f(x) := \bigsqcup_{i\in I} f_0(y_i)$. It is not clear whether f is well-defined, let alone continuous. So we must first prove that f is well-defined. To do this, we take any two directed nets $(y_i)_{i\in I}$ and $(z_k)_{k\in K}$ in X_0 such that $(y_i)_{i\in I} \to_X \bigsqcup_{i\in I} y_i = x$ and $(z_k)_{k\in K} \to_X \bigsqcup_{k\in K} z_k = x$. We shall show that

$$\bigsqcup_{i\in I} f_0(y_i) = \bigsqcup_{k\in K} f_0(z_k).$$

Now since $(y_i)_{i\in I} \to_X x$ and $x \sqsupseteq z_k$ for each $k \in K$, so $(y_i)_{i\in I} \to_X z_k$ for each $k \in K$ by the down-niceness of X. Then by the continuity of $f_0 : X_0 \longrightarrow Y$, we have $(f_0(y_i))_{i\in I} \to_Y f_0(z_k)$ for each $k \in K$ and hence the set $\{f_0(z_k) \mid k \in K\} \subseteq \lim_Y(f_0(y_i))_{i\in I}$. As the latter set is Scott-closed (since Y is a d-space), it follows that $\bigsqcup_{k\in K} f_0(z_k) \in \lim_Y(f_0(y_i))_{i\in I}$. So $(f_0(y_i))_{i\in I} \to_Y \bigsqcup_{k\in K} f_0(z_k)$. Also $\bigsqcup_{i\in I} f_0(y_i) \sqsupseteq f_0(y_i)$, the constant net $(\bigsqcup_{i\in I} f_0(y_i)) \to_Y \bigsqcup_{k\in K} f_0(z_k)$ by the up-niceness of Y. Similarly, the constant net $(\bigsqcup_{k\in K} f_0(z_k)) \to_Y \bigsqcup_{i\in I} f_0(y_i)$. Because Y is a T_0-space, we have that $\bigsqcup_{i\in I} f_0(y_i) = \bigsqcup_{k\in K} f_0(z_k)$ as desired. So f is well-defined.

We now show that f is continuous. Let $(x_k)_{k\in K} \to_X x$. We want to show that $(f(x_k))_{k\in K} \to_Y f(x)$. Since X_0 is a d-base for X, it follows that

1. for each $k \in K$, there exists a directed net $(y_j^k)_{j\in J(k)} \to_X \bigsqcup_{j\in J(k)} y_j^k = x_k$, and
2. there is a directed net $(y_j)_{j\in J} \to_X \bigsqcup_{j\in J} y_j = x$.

It is enough to show that $f(x_k) \sqsupseteq f_0(y_j)$ eventually for each $j \in J$, relying on the up-niceness of Y. For this purpose, we suppose that there exists $j_0 \in J$ such that for all $k' \in K$, there is $k \in K$ with $k \sqsupseteq k'$ such that $f(x_k) \not\sqsupseteq f_0(y_{j_0})$. Then $f_0(y_{j_0})$ belongs to the open set $Y\setminus \downarrow f(x_k) = Y\setminus \downarrow \bigsqcup_{j\in J(k)} f_0(y_j^k)$. Since X is an IL space, it follows that $(y_g^k(k))_{(k,g)\in K\times M} \to_X x$ where $M = \prod_{i\in I} J(i)$. Then by the down-niceness of X, this net will converge to y_{j_0}. Invoking the continuity of f_0, it follows that $(f_0(y_g^k(k)))_{(k,g)\in K\times M} \to_Y f_0(y_{j_0})$. Because $Y\setminus \downarrow \bigsqcup_{j\in J(k)} f_0(y_j^k)$ is an open set, one deduces that $f_0(y_{g(k)}^k)$ is in the same open set eventually, implying that $f_0(y^k_{g(k)}) \not\sqsubseteq \bigsqcup_{j\in J(k)} f_0(y^k_j)$ for some g, which is impossible. So the proof is completed.

Definition 6.6. *Given a space $(X,\rightarrow)$, a D-completion of $(X,\rightarrow)$ is a pair (Y,η) where Y is a d-space and $\eta : X\longrightarrow Y$ is a continuous mapping such that for each continuous mapping $f : X\longrightarrow Z$ into a d-space Z there is a unique continuous mapping $\hat{f} : Y\longrightarrow Z$ such that $f = \hat{f}\circ\eta$.*

By [2,6], every T_0 topological space has a D-completion. In [14], every poset has a dcpo-completion. Our main concern in this section is which net convergence spaces, apart from the topological ones, have a D-completions.

Theorem 6.7. *If X can be embedded into a IL d-space, then X has a D-completion.*

Proof. For the sake of convenience, we assume X is a subspace of an IL d-space Y. For each subspace A of Y, define

$$d(A) = \{y\in Y \mid \exists \text{ a directed family } x_i\subseteq A.(x_i)_{i\in I}\rightarrow_Y \bigsqcup_{i\in I} x_i = y\}.$$

Now for each ordinal α, we define a subset X^α by transfinite induction as follows:

$$X^1 = d(X),\ X^{\beta+1} = d(X^\beta) \text{ and } X^\alpha = \bigcup_{\gamma<\alpha} X^\gamma \text{ if } \alpha \text{ is a limit ordinal.}$$

By the usual ordinal reason, there is a smallest ordinal α satisfying the equation

$$X^{\alpha+1} = X^\alpha.$$

For any continuous mapping $g : X\longrightarrow Z$ from X into a d-space Z, there is a unique extension of g over X^1 by virtue of Lemma 6.7. By transfinite induction, it follows that f has a unique continuous extension over X^α.

Since every T_0 topological space is embedded into some Ω^M (c.f. Lemma II-3.4 of [3]), where Ω is the Sierpinski space and since Ω is clearly a d-space, so we have:

Corollary 6.8. *Every T_0 topological space has a D-completion which is also topological.*

7 Conclusion

In this paper, we have suggested a way of generalizing the concept of a d-space from topological spaces to net convergence spaces. In addition, we have obtained a sufficient condition that translates to a particular class of net convergence spaces that admits a d-completion to exist for each member of this class. Crucially, this class of net convergence spaces contains as members all T_0 topological spaces, and thus we have generalized existing results concerning d-completions of T_0-spaces and dcpo-completions of posets. However, the authors have not justified at the moment of writing whether such a class of net convergence spaces *properly* contains the category of T_0 spaces. Thus any future works leading from here must crucially include the construction of an example of a non-topological net convergence space which can be embedded into an d-space which has all iterated limits.

References

1. Abramsky, S., Jung, A.: Domain Theory. Handbook of Logic in Computer Science, vol. 3. Clarendon Press, Oxford (1994)
2. Ershov, Y.L.: On d-spaces. Theoretical Computer Science 224, 59–72 (1999)
3. Gierz, G., Hofmann, K.H., Keimel, K., Lawson, J.D., Mislove, M.W., Scott, D.S.: Continuous Lattices and Domains. Encyclopedia of Math- ematics and its Applications, vol. 93. Cambridge University Press, Cambridge (2003)
4. Heckmann, R.: A non-topological view of dcpos as convergence spaces. Theoretical Computer Science 305, 159 (2003)
5. Johnstone, P.T.: Stone Spaces. Cambridge Studies in Advanced Mathematics, vol. 3. Cambridge University Press, Cambridge (1982)
6. Keimel, K., Lawson, J.D.: D-completions and d-topology. Annals of Pure and Applied Logic 159(3), 292–306 (2009)
7. Kelley, J.L.: General Topology. Graduate Texts in Mathematics, vol. 27. Springer, Heidelberg (1975)
8. Markowsky, G.: Chain-complete p.o. sets and directed sets with applications. Algebra Universalis 6, 53–68 (1976)
9. Moore, E.H., Smith, H.L.: A general theory of limits. American Journal of Mathematics 44, 102–121 (1922)
10. Murdeshwar, M.G.: General Topology. New Age International (P) Limited (1999)
11. Preuss, G.: Theory of Topological Structures. D. Reidel Publishing Company, Dordrecht (1988)
12. Wyler, O.: Dedekind complete posets and scott topologies. In: Banaschewski, B., Hokmann, R.-E. (eds.) Continuous Lattices. Lecture Notes in Mathematics, vol. 871, pp. 383–389 (1981)
13. Zhao, B., Zhao, D.: The lim-inf convergence on partially ordered sets. J. Mathematical Analysis and its Applications 309, 701–708 (2005)
14. Zhao, D., Fan, T.: Dcpo-completion of posets. Theoretical Computer Science 411, 2167–2173 (2010)

Modal R_0-Algebra-Valued Modal Logic System M$\mathcal{L}^*$

Hui-Xian Shi and Guo-Jun Wang

Institute of Mathematics, Shaanxi Normal University, Xi'an 710062, China
e-mail: rubyshi@163.com, gjwang@snnu.edu.cn

Abstract. Based on the concept of the R_0-algebra, the present paper introduces the definition of the modal R_0-algebra by adding a new unary operator $\Box$, corresponding to modalities of the modal logic. By use of modal R_0-algebras, semantic and syntactic frameworks are constructed, respectively, for logic system M$\mathcal{L}^*$, the modal R_0-algebra-valued modal logic system. It is pointed out that the semantics of system M$\mathcal{L}^*$ generalizes the semantics of both the classical modal logic and the $[0, 1]$-valued modal logic. The main result of the paper is the completeness theorem of system M$\mathcal{L}^*$.

Keywords: Modal R_0-algebra, Modal Logic System, Modal Model, Validity, Completeness.

1 Introduction

The formalism of modal logic and its proof technique, one of the most efficient tools for knowledge representation and reasoning about discrete dynamic systems, plays a significant role in artificial intelligence and computer science [1, 2]. Several systems with various kinds of modal operators have been constructed [3, 4, 5, 6], and corresponding applications have been found and researched in fuzzy concept analysis and other important realms [7, 8]. Meanwhile, by taking into account the numerical calculation for mathematical logic in order to grade the concepts of truth values for formulas, [9, 10, 11, 12] have proposed the theory of quantitative logic, and this quantitative approach has been considered in the classical modal logic [12]. Besides, by generalizing the classical Kripke models, the semantics of $[0, 1]$-valued modal logic has been introduced [13, 14, 15].

It is worth noting that there are limitations in [4], although it has proved several completeness theorems. There the Kripke models have the form $K = (W, e, A)$ where A is required as a chain, and the logic is only available for the generalization of the modal logic system $S5$, for the sake of corresponding modalities $\Box$ and $\Diamond$ in the modal logic to universal quantifier

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 111–120.
springerlink.com

$\forall$ and existential quantifier $\exists$, respectively, in the predicate logic containing only one variable. In order to modify these limitations, the present paper intends to construct a Kripke model not limited to totally-ordered ones and a logic system available not only for the generalization of $S5$. Based on the concept of the R_0-algebra, the definition of the modal R_0-algebra (briefly, MR_0-algebra) is introduced by adding a new unary operator $\Box$, corresponding to modalities of the modal logic. Considering complete MR_0-algebras as valuation fields, the paper proposes the concept of the MR_0-modal model and constructs the semantics in the class of MR_0-modal models. It is pointed out that this semantics generalizes the semantics of both the classical modal logic and the $[0,1]$-valued modal logic. Meanwhile, the MR_0-algebra-valued modal logic system M$\mathcal{L}^*$ is constructed, and the completeness theorem of system M$\mathcal{L}^*$ is obtained, i.e., a modal formula is a theorem in M$\mathcal{L}^*$ if and only if it is MR_0-valid.

2 Preliminaries

The language of the classical modal logic [1, 3] is generated by the form below:

$$\varphi := p \mid \bot \mid \neg\varphi \mid \varphi_1 \vee \varphi_2 \mid \Diamond\varphi, \quad p \in \Phi,$$

where Φ is the set of propositional variables, $\bot$ denotes the contradiction, and

$\Box\varphi$ is $\neg\Diamond\neg\varphi$,
$\varphi \wedge \psi$ is $\neg(\neg\varphi \vee \neg\psi)$,
$\varphi \to \psi$ is $\neg\varphi \vee \psi$.

The set of all classical modal formulas is denoted by Form$(\Diamond, \Phi)$.

Definition 1. *([3]) A Kripke model for the classical modal logic (briefly, classical model) is a triple $\boldsymbol{M} = (W, R, V)$, where W is a nonempty set of possible worlds, $R \subset W \times W$ is a binary relation on W, $V : \Phi \longrightarrow \mathcal{P}(W)$ is a mapping, where $\mathcal{P}(W)$ is the power set of W.*

Let $\varphi \in$Form$(\Diamond, \Phi)$, $w \in W$. The world w satisfying φ, denoted by $\boldsymbol{M}, w \models \varphi$, can be recursively defined as follows:

i) $\boldsymbol{M}, w \models p$ if and only if $w \in V(p)$, $p \in \Phi$.
ii) $\boldsymbol{M}, w \models \bot$ never holds.
iii) $\boldsymbol{M}, w \models \neg\varphi$ if and only if $\boldsymbol{M}, w \models \varphi$ does not hold.
iv) $\boldsymbol{M}, w \models \varphi \vee \psi$ if and only if $\boldsymbol{M}, w \models \varphi$ or $\boldsymbol{M}, w \models \psi$.
v) $\boldsymbol{M}, w \models \Diamond\varphi$ if and only if $\exists u \in W, (w, u) \in R$ s.t. $\boldsymbol{M}, u \models \varphi$.

Definition 2. *([3]) Let $\varphi \in$Form$(\Diamond, \Phi)$. Say that φ is valid if $\boldsymbol{M}, w \models \varphi$ holds for every classical model $\boldsymbol{M} = (W, R, V)$ and for every world $w \in W$.*

Proposition 1. *([3]) Let $\boldsymbol{M} = (W, R, V)$ be a classical model. And define*

$$V(\varphi) = \{w \in W \mid \boldsymbol{M}, w \models \varphi\}, \quad \varphi \in Form(\Diamond, \Phi),$$

then

i) $V(\neg\varphi) = W - V(\varphi)$;

ii) $V(\varphi \vee \psi) = V(\varphi) \cup V(\psi)$;

iii) $V(\varphi \wedge \psi) = V(\varphi) \cap V(\psi)$;

iv) $V(\varphi \to \psi) = (W - V(\varphi)) \cup V(\psi)$;

v) $V(\Diamond\varphi) = \{w \in W | \ R[w] \cap V(\varphi) \neq \emptyset\}$,

where $R[w] = \{u \in W | \ (w, u) \in R\}$.

3 Modal R_0-Algebras

Definition 3. *([9]) Let* L *be an algebra of type* $(\vee, \wedge, \ ', \to)$*, where* $'$ *is a unary operator,* $\vee, \wedge$ *and* $\to$ *are binary operators.* $(L, \vee, \wedge, \ ', \to)$ *is called an* R_0*-algebra if there is a partial order* $\leq$ *such that* $(L, \leq, 1)$ *is a bounded distributive lattice with the greatest element 1,* $\vee$ *and* $\wedge$ *are the supremum and infimum operators, respectively, with respect to* $\leq$*,* $'$ *is an order-reversing involution, and the following conditions hold for every* $a, b, c \in L$*:*

(m1) $a' \to b' = b \to a$.

(m2) $1 \to a = a, \quad a \to a = 1$.

(m3) $b \to c \leq (a \to b) \to (a \to c)$.

(m4) $a \to (b \to c) = b \to (a \to c)$.

(m5) $a \to b \vee c = (a \to b) \vee (a \to c), \quad a \to b \wedge c = (a \to b) \wedge (a \to c)$.

(m6) $(a \to b) \vee ((a \to b) \to a' \vee b) = 1$.

Definition 4. *Let* $L = L(\vee, \wedge, \ ', \to)$ *be an* R_0*-algebra and* $\Box$ *be a unary operator on* L*.* $(L, \vee, \wedge, \ ', \to, \Box)$ *is called a modal* R_0*-algebra,* MR_0*-algebra in brief, if the following conditions hold for every* $a, b \in L$*:*

(m7) $\Box(a \to b) \leq \Box a \to \Box b$.

(m8) $\Box 1 = 1$.

Example 1. (i) Let $L = (\{0, 1\}, \vee, \wedge, \ ')$ be a $\{0, 1\}$-Boole algebra. Then L is obviously a bounded distributive lattice. Define

$$a \to b = a' \vee b, \quad a, b \in L,$$

then it can be proved that $L = (\{0, 1\}, \vee, \wedge, \ ', \to)$ satisfies the conditions (m1)-(m6) above and consequently becomes an R_0-algebra. Furthermore, if we define a unary operator $\Box$ on L satisfying the conditions (m7) and (m8), e.g., $\Box 0 = \Box 1 = 1$, then $L = (\{0, 1\}, \vee, \wedge, \ ', \to, \Box)$ will become an MR_0-algebra.

(ii) Let $L = [0, 1]$ and $\vee, \wedge$ be the supremum and infimum operators, respectively, with respect to the usual order on the real line. Then it is obvious that L is a bounded distributive lattice. Define

$$a' = 1 - a,$$

$$a \to b = \begin{cases} 1, & a \leq b, \\ (1 - a) \vee b, & a > b, \end{cases}$$

where $a, b \in L$, then $'$ becomes an order-reversing involution on L, and $\to$ the R_0-implication operator [9] on L. It can be proved that $L = ([0,1], \vee, \wedge, \ ', \to)$ satisfies the conditions (m1)-(m6) and consequently becomes an R_0-algebra. Furthermore, if we define a unary operator $\Box$ on L satisfying the conditions (m7) and (m8), e.g.,

$$\Box a = \begin{cases} 0, & a \leq \frac{1}{2}, \\ a, & a > \frac{1}{2}, \end{cases}$$

then $L = ([0,1], \vee, \wedge, \ ', \to, \Box)$ will become an MR_0-algebra.

Proposition 2. *Let $L = L(\vee, \wedge, \ ', \to, \Box)$ be a complete MR_0-algebra. Then the following conditions hold for every $a, b, c \in L$ and $Q \subseteq L$:*

(i) $a' = a \to 0$.
(ii) $a \to b = 1$ *if and only if* $a \leq b$.
(iii) $a \to (b \to a \wedge b) = 1$.
(iv) If $b \leq c$, *then* $a \to b \leq a \to c$; *If* $a \leq b$, *then* $b \to c \leq a \to c$.
(v) $a \vee b \to c = (a \to c) \wedge (b \to c)$, $a \wedge b \to c = (a \to c) \vee (b \to c)$.
(vi) $(\vee_{a \in Q}\ a)' = \wedge_{a \in Q}\ a'$, $(\wedge_{a \in Q}\ a)' = \vee_{a \in Q}\ a'$.

4 The MR_0-Algebra-Valued Modal Logic System M$\mathcal{L}^*$

4.1 Semantics of System M$\mathcal{L}^*$

The language of the MR_0-algebra-valued modal logic system M$\mathcal{L}^*$ is generated by the form

$$\varphi := p \mid \perp \mid \neg\varphi \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \to \varphi_2 \mid \Diamond\varphi, \quad p \in \Phi,$$

where Φ is the set of propositional variables, $\perp$ denotes the contradiction, and

$\Box\varphi$ is $\neg\Diamond\neg\varphi$,
$\varphi \wedge \psi$ is $\neg(\neg\varphi \vee \neg\psi)$.
The set of all modal formulas of M$\mathcal{L}^*$ is denoted by F(Φ).

Note that, F(Φ) is different from Form($\Diamond, \Phi$), the set of all classical modal formulas, since $\varphi \to \psi$ and $\neg\varphi \vee \psi$ are two distinct formulas in F(Φ). However, these two formulas are equal to each other in Form($\Diamond, \Phi$).

Definition 5. *A Kripke model for the MR_0-algebra-valued modal logic is a quadruple $\mathcal{M}= (W, R, e, L)$, where W is a nonempty set of possible worlds, $L = L(\vee, \wedge, \ ', \to, \Box)$ is a complete MR_0-algebra, $R : W \times W \longrightarrow L$ is an L-valued binary relation on W, and e is a mapping $e : W \times \Phi \longrightarrow L$ assigning to each variable in Φ a truth value belonging to L. L is said to be a valuation field, and e a valuation mapping. Besides, a Kripke model for the MR_0-algebra-valued modal logic can be briefly called an MR_0-modal model.*

Definition 6. *Let $\mathcal{M}$= (W, R, e, L) be an MR_0-modal model. The valuation mapping e can be uniquely extended to a mapping $\bar{e} : W \times F(\Phi) \longrightarrow L$ satisfying that:*

$$\bar{e}(w, \neg\varphi) = \bar{e}(w, \varphi)',$$
$$\bar{e}(w, \varphi \vee \psi) = \bar{e}(w, \varphi) \vee \bar{e}(w, \psi),$$
$$\bar{e}(w, \varphi \to \psi) = \bar{e}(w, \varphi) \to \bar{e}(w, \psi),$$
$$\bar{e}(w, \Diamond\varphi) = \vee\{\bar{e}(u, \varphi)|\ R(w, u) \neq 0, u \in W\},$$

where $w \in W$, $\varphi, \psi \in F(\Phi)$.

Although the mappings e and $\bar{e}$ are different, there will be no confusion between them, and so we will use the same notation e for both in the following.

It is not difficult to infer from Proposition 2 and Definition 6 that.

Corollary 1. *Let $\mathcal{M}$= (W, R, e, L) be an MR_0-modal model, $w \in W$, $\varphi \in F(\Phi)$. Then*

$$e(w, \varphi \wedge \psi) = e(w, \varphi) \wedge e(w, \psi),$$
$$e(w, \Box\varphi) = \wedge\{e(u, \varphi)|\ R(w, u) \neq 0, u \in W\}.$$

Example 2. Suppose that the valuation field $L = (\{0, 1\}, \vee, \wedge, \ ', \to, \Box)$ is the MR_0-algebra defined in Example 1(i), then L is obviously complete. Let $W \neq \emptyset$ be a set, $R : W \times W \longrightarrow \{0, 1\}$, and $e : W \times \Phi \longrightarrow \{0, 1\}$, then $\mathcal{M}$= $(W, R, e, \{0, 1\})$ becomes an MR_0-modal model. We omit $\{0, 1\}$ in the model $\mathcal{M}$, and denote it by $\mathcal{M}$= (W, R, e). Note that, here the relation R has become a classical binary relation on W, i.e., $R \subset W \times W$. Therefore, the set W of possible worlds and binary relation R in the model $\mathcal{M}$= (W, R, e) are same with the ones in the classical models. The difference lies in the valuation mappings between these two kinds of models, since the valuation mapping in the classical models is the mapping $V : \Phi \longrightarrow \mathcal{P}(W)$.

Meanwhile, if $\boldsymbol{M} = (W, R, V)$ is a classical model, then the mapping $V : \Phi \longrightarrow \mathcal{P}(W)$ can induce a mapping $V^* : W \times \Phi \longrightarrow \{0, 1\}$ by:

$$V^*(w, p) = \begin{cases} 1, & w \in V(p), \\ 0, & w \notin V(p). \end{cases}$$

It can be proved that the mappings V and V^*, after extending following Proposition 1 and Definition 6, respectively, satisfy that

$$V^*(w, \varphi) = 1 \ \text{ iff } \ w \in V(\varphi) \ \text{ iff } \ \boldsymbol{M}, w \models \varphi, \quad w \in W, \varphi \in \text{Form}(\Diamond, \Phi).$$

Therefore, despite of the difference between their definitions, the mappings V and V^* can be considered to be same in nature.

In addition, it can be seen that the valuation mapping $e : W \times \Phi \longrightarrow \{0, 1\}$ in the MR_0-modal model above is actually in accordance with the mapping $V^* : W \times \Phi \longrightarrow \{0, 1\}$, as well as the mapping V, consequently. In this sense, we can see that the MR_0-modal model $\mathcal{M}$= (W, R, e) above actually becomes a classical model. As a result, the semantics of the MR_0-algebra-valued modal logic is the generalization of the semantics of the classical modal logic.

Example 3. Suppose that the valuation field $L = ([0,1], \vee, \wedge, \ ', \rightarrow, \Box)$ is the MR_0-algebra defined in Example 1(ii), then L is obviously complete. Let $W \neq \emptyset$ be a set, $R : W \times W \longrightarrow [0,1]$ be a fuzzy binary relation on W, and $e : W \times \Phi \longrightarrow [0,1]$, then $\mathcal{M}= (W, R, e, [0,1])$ becomes an MR_0-modal model.

Meanwhile, by considering the unit interval $[0,1]$ as a valuation field, [13] has proposed the semantics of the $[0,1]$-valued modal logic(also called fuzzy modal propositional logic in [13]), whose model is a triple $\boldsymbol{k} = (U, R, I)$, where U is a nonempty set of possible worlds, $R : U \times U \longrightarrow [0,1]$ is a fuzzy binary relation on U, $I : U \times S \longrightarrow [0,1]$ is a mapping, where S is the set of propositional variables. Note that, this model is actually in accordance with the MR_0-modal model $\mathcal{M}= (W, R, e, [0,1])$ above, without considering the difference in their notations. As a result, the semantics of the MR_0-algebra-valued modal logic is also the generalization of the semantics of the $[0,1]$-valued modal logic.

Definition 7. *Let $\varphi \in F(\Phi)$. Say that φ is MR_0-valid if $e(w, \varphi) = 1$ holds for every MR_0-modal model $\mathcal{M}= (W, R, e, L)$ and for every world $w \in W$.*

4.2 Syntactics of System M$\mathcal{L}^*$

The axioms and inference rules for the system M$\mathcal{L}^*$ are as follows:

- Axioms:
  - (M1) $\varphi \rightarrow (\psi \rightarrow \varphi \wedge \psi)$
  - (M2) $(\neg\varphi \rightarrow \neg\psi) \rightarrow (\psi \rightarrow \varphi)$
  - (M3) $(\varphi \rightarrow (\psi \rightarrow \gamma)) \rightarrow (\psi \rightarrow (\varphi \rightarrow \gamma))$
  - (M4) $(\psi \rightarrow \gamma) \rightarrow ((\varphi \rightarrow \psi) \rightarrow (\varphi \rightarrow \gamma))$
  - (M5) $\varphi \rightarrow \neg\neg\varphi$
  - (M6) $\varphi \rightarrow \varphi \vee \psi$
  - (M7) $\varphi \vee \psi \rightarrow \psi \vee \varphi$
  - (M8) $(\varphi \rightarrow \gamma) \wedge (\psi \rightarrow \gamma) \rightarrow (\varphi \vee \psi \rightarrow \gamma)$
  - (M9) $(\varphi \wedge \psi \rightarrow \gamma) \rightarrow (\varphi \rightarrow \gamma) \vee (\psi \rightarrow \gamma)$
  - (M10) $(\varphi \rightarrow \psi) \vee ((\varphi \rightarrow \psi) \rightarrow \neg\varphi \vee \psi)$
  - (K) $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$
  - ($\bot$) $\bot \rightarrow \Diamond\bot$
- Inference rules:
  - Modus Ponens: from φ and $\varphi \rightarrow \psi$ infer ψ.
  - Necessitation: from φ infer $\Box\varphi$.

Remark 1. Note that, axioms (M1)-(M10) of M$\mathcal{L}^*$ are actually the substitution instances [3] of axioms ($\mathcal{L}^*$1)-($\mathcal{L}^*$10) of the system $\mathcal{L}^*$ proposed in [9]. Consequently, it can be inferred that a theorem in $\mathcal{L}^*$ is also a theorem in M$\mathcal{L}^*$. However, the two axioms (K) and ($\bot$), related to the modality $\Box$, of M$\mathcal{L}^*$ are not necessary to be theorems of $\mathcal{L}^*$. As a result, the MR_0-algebra-valued modal logic system above is called the modal $\mathcal{L}^*$ system, briefly denoted by M$\mathcal{L}^*$.

Meanwhile, it is indicated in [3] that the set of axioms for the classical modal logic system $\boldsymbol{K}$ includes the formula (K) and all the modal tautologies(see [3]), i.e., the substitution instances of theorems from the classical two-valued propositional logic system $\boldsymbol{L}$ [9]. It can be proved that the axioms (M1)-(M9) in M$\mathcal{L}^*$ are all modal tautologies in $\boldsymbol{K}$, and consequently are theorems of $\boldsymbol{K}$. However, the axiom $(\varphi \to (\psi \to \gamma)) \to ((\varphi \to \psi) \to (\varphi \to \gamma))$ of $\boldsymbol{K}$, also the substitution instance of the axiom (L2) from $\boldsymbol{L}$, is not necessary to be a theorem of M$\mathcal{L}^*$.

Proposition 3. *In the system M$\mathcal{L}^*$,*

(i) Hypothetical Syllogism holds, i.e., from $\varphi \to \psi$ and $\psi \to \gamma$ infer $\varphi \to \gamma$.

(ii) $\varphi \to \varphi$ and $(\varphi \to \psi) \to (\neg\psi \to \neg\varphi)$ are theorems.

(iii) If φ and ψ are theorems, then $\varphi \wedge \psi$ is also a theorem.

(iv) If $\varphi \to \psi$ and $\gamma \to \chi$ are theorems, then $\varphi \vee \gamma \to \psi \vee \chi$ is also a theorem;

If $\varphi \to \gamma$ and $\psi \to \gamma$ are theorems, then $\varphi \vee \psi \to \gamma$ is also a theorem.

Proposition 4. *(Soundness theorem of M$\mathcal{L}^*$) The theorems in M$\mathcal{L}^*$ are all MR_0-valid.*

Proof. First of all, we prove that all the inference rules of M$\mathcal{L}^*$ preserve MR_0-validity of formulas:

Suppose that φ and $\varphi \to \psi$ are MR_0-valid, then $e(w, \varphi) = e(w, \varphi \to \psi) = 1$ holds for every MR_0-modal model $\mathcal{M} = (W, R, e, L)$ and for every $w \in W$. Therefore $e(w, \varphi \to \psi) = e(w, \varphi) \to e(w, \psi) = 1$, and by Proposition 2(ii) we can infer that $1 = e(w, \varphi) \leq e(w, \psi)$, i.e., $e(w, \psi) = 1$. Since the model $\mathcal{M}$ and the world w are arbitrary, then ψ is MR_0-valid.

Suppose that φ is MR_0-valid, then $e(w, \varphi) = 1$ holds for every MR_0-modal model $\mathcal{M}= (W, R, e, L)$ and for every $w \in W$. Therefore, in the condition that $\triangle_w = \{u \in W |\ R(w, u) \neq 0\} \neq \emptyset$, we can obtain $e(u, \varphi) = 1$ holds for every $u \in \triangle_w$, and consequently $e(w, \Box\varphi) = \wedge_{u \in \triangle_w} e(u, \varphi) = 1$. If $\triangle_w = \emptyset$, then it is obvious that $e(w, \Box\varphi) = \wedge\emptyset = 1$ holds. Since the model $\mathcal{M}$ and the world w are arbitrary, then $\Box\varphi$ is MR_0-valid.

From above, we can conclude that MP and Necessitation all preserve MR_0-validity. In addition, it is not difficult to prove by Proposition 2 that all the axioms in M$\mathcal{L}^*$ are MR_0-valid. Since all the theorems of M$\mathcal{L}^*$ can be deducted from axioms by inference rules within finite steps, hence they are MR_0-valid. The proof is completed.

4.3 Completeness of System M$\mathcal{L}^*$

In M$\mathcal{L}^*$, the provable equivalence relation on F(Φ), denoted by $\sim$, can be defined in the usual way, i.e., $\varphi \sim \psi$ if and only if $\varphi \to \psi$ and $\psi \to \varphi$ are theorems of M$\mathcal{L}^*$.

Proposition 5. *The provable equivalence relation $\sim$ is a congruence relation* [16] *on F(Φ) of type* $(\neg, \vee, \to, \Box)$.

Proof. Firstly, it can be easily proved by Proposition 3(i)(ii) that $\sim$ is an equivalence relation on $\mathrm{F}(\Phi)$, and we further prove that $\sim$ is also a congruence relation on $\mathrm{F}(\Phi)$ of type $(\neg, \vee, \rightarrow, \Box)$ as follows:

(i) Let $\varphi \sim \psi$, then $\varphi \rightarrow \psi$ is a theorem. It follows from Proposition 3(ii) and MP that $\neg\psi \rightarrow \neg\varphi$ is a theorem. The other direction implication is given by a similar argument. Therefore $\neg\varphi \sim \neg\psi$.

(ii) Let $\varphi \sim \psi$ and $\gamma \sim \chi$, then $\varphi \rightarrow \psi$ and $\gamma \rightarrow \chi$ are theorems. By Proposition 3(iv), $\varphi \vee \gamma \rightarrow \psi \vee \chi$ is a theorem. Similarly, it can be proved that $\psi \vee \chi \rightarrow \varphi \vee \gamma$ is also a theorem. Hence $\varphi \vee \gamma \sim \psi \vee \chi$.

(iii) Let $\varphi \sim \psi$ and $\gamma \sim \chi$, then $\gamma \rightarrow \chi$ is a theorem. It follows from (M4) and MP that $(\varphi \rightarrow \gamma) \rightarrow (\varphi \rightarrow \chi)$ is a theorem. Similarly, we obtain $(\neg\chi \rightarrow \neg\varphi) \rightarrow (\neg\chi \rightarrow \neg\psi)$ is a theorem from $\neg\varphi \sim \neg\psi$. Then it follows from (M2) and Proposition 3(ii) by HS twice that $(\varphi \rightarrow \chi) \rightarrow (\psi \rightarrow \chi)$ is a theorem. Again by HS, we obtain $(\varphi \rightarrow \gamma) \rightarrow (\psi \rightarrow \chi)$ is a theorem. The other direction implication is given by a similar argument. Hence, $\varphi \rightarrow \gamma \sim \psi \rightarrow \chi$.

(iv) Let $\varphi \sim \psi$, then $\varphi \rightarrow \psi$ is a theorem. It follows from Necessitation that $\Box(\varphi \rightarrow \psi)$ is a theorem. Then by (K) and MP, we obtain $\Box\varphi \rightarrow \Box\psi$ is a theorem. Similarly, we can prove the other direction implication. Therefore, $\Box\varphi \sim \Box\psi$. The proof is completed.

Since the provable equivalence relation $\sim$ is a congruence relation on $\mathrm{F}(\Phi)$ of type $(\neg, \vee, \rightarrow, \Box)$, then we obtain a quotient class of $\mathrm{F}(\Phi)$ by $\sim$, denoted by $\mathcal{F}$, and it can be inferred that $\mathcal{F}$, with the operations inherited from $\mathrm{F}(\Phi)$, is an algebra of type $(\neg, \vee, \rightarrow, \Box)$. The elements of $\mathcal{F}$ are denoted by $[\varphi]$ $(\varphi \in \mathrm{F}(\Phi))$, where $[\varphi] = \{\psi \in \mathrm{F}(\Phi) | \ \varphi \sim \psi\}$.

Proposition 6. *The quotient algebra $\mathcal{F} = F(\Phi)/\sim$ is an MR_0-algebra, in which the partial order $\leq$ is defined by*

$$[\varphi] \leq [\psi] \ \ \textit{if and only if} \ \ \varphi \rightarrow \psi \textit{ is a theorem}, \tag{1}$$

and the operators $\neg, \vee, \rightarrow, \Box$ on $\mathcal{F}$ are defined by

$$\neg[\varphi] = [\neg\varphi], \quad [\varphi] \vee [\psi] = [\varphi \vee \psi], \quad [\varphi] \rightarrow [\psi] = [\varphi \rightarrow \psi], \quad \Box[\varphi] = [\Box\varphi]. \tag{2}$$

Proof. It is easy to prove that $\leq$ defined in (1) is well-defined and indeed a partial order on $\mathcal{F}$. Since the provable equivalence relation $\sim$ is a congruence relation on $\mathrm{F}(\Phi)$ of type $(\neg, \vee, \rightarrow, \Box)$, then the operators $\neg, \rightarrow, \Box$ defined in (2) are well-defined, and it is trifles to prove that $\neg$ is an order-reversing involution on $\mathcal{F}$.

Besides, since $\varphi \rightarrow \varphi\vee\psi$ and $\psi \rightarrow \varphi\vee\psi$ are theorems, we obtain that $[\varphi\vee\psi]$ is an upper bound of $\{[\varphi], [\psi]\}$ w.r.t. $\leq$. Let $[\gamma]$ be an arbitrary upper bound of $\{[\varphi], [\psi]\}$, then $\varphi \rightarrow \gamma$ and $\psi \rightarrow \gamma$ are theorems by (1), and consequently $\varphi \vee \psi \rightarrow \gamma$ is a theorem by Proposition 3(iv). Hence $[\varphi \vee \psi] \leq [\gamma]$, which indicates that $[\varphi \vee \psi]$ is just the supremum of $\{[\varphi], [\psi]\}$ w.r.t. $\leq$. Thus the operator $\vee$ defined in (2) is also well-defined.

Let $\top$ be a theorem in F(Φ), then it can be proved that $\varphi \to \top$ is a theorem for every $\varphi \in$F(Φ), and hence $[\varphi] \leq [\top]$. This shows that $[\top]$ is the greatest element of $\mathcal{F}$, i.e., 1. Similarly, $[\bot]$ is the smallest element of $\mathcal{F}$, i.e., 0.

Lastly, it follows from Proposition 3 and (2) that the quotient algebra $\mathcal{F}$ satisfies the conditions (m1)-(m8) in Definition 3 and 4, and consequently is an MR_0-algebra. The proof is completed.

Theorem 1. *(Completeness theorem of M$\mathcal{L}^*$) Let $\varphi \in F(\Phi)$. φ is a theorem of M$\mathcal{L}^*$ if and only if φ is MR_0-valid.*

Proof. By Proposition 4, it suffices to prove that all MR_0-valid formulas are theorems of M$\mathcal{L}^*$.

Assume that φ is MR_0-valid. Then an MR_0-modal model $\mathcal{M}=(W,R,e,L)$ can be defined as follows:

Let $W \neq \emptyset$ be a set, the valuation field L be the quotient algebra $\mathcal{F}=$F$(\Phi)/\sim$, $R: W \times W \longrightarrow \mathcal{F}$ be an $\mathcal{F}$-valued binary relation on W, and the valuation mapping $e: W \times \Phi \longrightarrow \mathcal{F}$ be defined by

$$e(w,p) = [p], \quad p \in \Phi, w \in W.$$

It follows from (2) that the valuation mapping e, after extending following Definition 6, satisfies that

$$e(w,\varphi) = [\varphi], \quad \varphi \in \mathrm{F}(\Phi), w \in W.$$

Since φ is MR_0-valid, then for the MR_0-modal model $\mathcal{M}=(W,R,e,L)$ defined above and $\forall w \in W$, $e(w,\varphi) = [\varphi] = 1 = [\top]$ holds, where $\top$ is a theorem of M$\mathcal{L}^*$. As a result, φ is a theorem of M$\mathcal{L}^*$, and the proof is complete.

5 Conclusions

In the present paper, the definition of the modal R_0-algebra (briefly, MR_0-algebra) is introduced first of all, based on the concept of the R_0-algebra, by adding a new unary operator $\Box$, corresponding to modalities of the modal logic. Secondly, by considering the complete MR_0-algebra as a valuation field, the paper proposes the concept of the MR_0-modal model and constructs the semantics in the class of MR_0-modal models. It is pointed out that this semantics generalizes the semantics of both the classical modal logic and the $[0,1]$-valued modal logic. Lastly, the MR_0-algebra-valued modal logic system M$\mathcal{L}^*$ is constructed, and the completeness theorem of system M$\mathcal{L}^*$ is obtained. Note that, various modal models can be defined as the valuation fields change, and different semantic and syntactic frameworks can be constructed in correspondence, which will be investigated in a forthcoming paper.

Acknowledgements. This work described here is partially supported by the grants from the National Natural Science Foundation of China (NSFC No.10771129).

References

1. Blackburn, P., Rijke, M., de, V.Y.: Modal Logic. Cambridge University Press, New York (2001)
2. Wang, Y.T.: Introduction to Modern Logics. China Renmin University Press, Beijing (1987) (in Chinese)
3. Wang, G.J.: Non-classical Mathematical Logic and Approximate Reasoning, 2nd edn. Science Press, Beijing (2008) (in Chinese)
4. Hájek, P.: On fuzzy modal logics $S5(\mathcal{C})$. Fuzzy Sets and Systems (2009), doi:10.1016/j.fss.2009.11.011
5. Mironov, A.M.: Fuzzy modal logics. Journal of Mathematical Sciences 128(6), 3461–3483 (2005)
6. Ying, M.S.: On standard models of fuzzy modal logics. Fuzzy Sets and Systems 26(3), 357–363 (1988)
7. Huynh, V.N., Nakamori, Y., Ho, T.B., Resconi, G.: A context model for fuzzy concept analysis based upon modal logic. Information Sciences 160, 111–129 (2004)
8. Zhang, Z.Y., Sui, Y.F., Cao, C.G., Wu, G.H.: A formal fuzzy reasoning system and reasoning mechanism based on propositional modal logic. Theoretical Computer Science 368, 149–160 (2006)
9. Wang, G.J., Zhou, H.J.: Introduction to mathematical logic and resolution principle. Science Press/Alpha Science International Limited, Beijing, Oxford (2009)
10. Wang, G.J.: A universal theory of measure and integral on valuation spaces with respect to diverse implication operators. Science in China (Series E) 43(6), 586–594 (2000)
11. Wang, G.J., Zhou, H.J.: Quantitative logic. Information Sciences 179, 226–247 (2009)
12. Wang, G.J., Duan, Q.L.: Theory of (n) truth degrees of formulas in modal logic and a consistency theorem. Science in China (Series F) 52(1), 70–83 (2009)
13. Chen, T.Y., Wang, D.G.: The semantics of fuzzy modal propositional logic. Journal of Liaoning Normal University (Natural Science Edition) 26(4), 341–343 (2003) (in Chinese)
14. Wang, D.G., Gu, Y.D., Li, H.X.: Generalized tautology in fuzzy modal propositional logic. Acta Electronica Sinica 35(2), 261–264 (2003) (in Chinese)
15. Hu, M.D., Wang, G.J.: Tautologies and quasi-tautologies in fuzzy modal logic. Acta Electronica Sinica 37(11), 2484–2488 (2009) (in Chinese)
16. Davey, B.A., Priestley, H.A.: Introduction to Lattices and Order. Cambridge University Press, New York (1990)

Deduction Theorem for Symmetric Cirquent Calculus

Wen-yan Xu and San-yang Liu

Department of Applied Mathematics, Xidian University, Xi'an 710071 China
e-mail: xwykiki@126.com, liusanyang@126.com

Abstract. Cirquent calculus is a recent approach to proof theory, whose characteristic feature is being based on circuit-style structures (called cirquents) instead of the traditional formulas or sequents. In this paper we prove the deduction theorem for the symmetric version of cirquent calculus, and show that the derivation in the deduction theorem will be at most polynomially longer than the proof of implication, and vice versa.

Keywords: Proof Theory, Cirquent Calculus, Deduction Theorem.

1 Introduction

Cirquent calculus, introduced by G.Japaridze [1, 10], is a new proof-theoretic and semantic framework, whose characteristic feature is being based on circuit-style structures (called cirquents), as opposed to the traditional formulas or sequents. Among its advantages are higher efficiency and greater expressiveness. Classical logic is just a special, conservative fragment of the logic in the form of cirquent calculus, obtained by considering only *circuits* (different from the one in [7] where the "circuit" refers to the structure of deductions), i.e., cirquents where multiple identical-label ports are not allowed. Actually, [1] borrows many ideas and techniques from the calculus of structures [2, 3, 8, 9], especially deep inference that modify cirquents at any level rather than only around the root as is the case in sequent calculus. It elaborated a deep cirquent calculus system CL8 for computability logic [4, 5, 6] and discussed some possible variations of CL8, including a symmetric system CL8S. But the properties of CL8S were not dissected.

The rules of CL8 consist of restructuring rules and main rules. The main rules include coupling, weakening, and pulldown. The restructuring rules include deepening (flattening), globalization (localization), and lengthening (shortening), where each rule comes in two versions: one for $\bullet$ and one for $\circ$. So, altogether there are 12 restructuring rules and 3 main rules. System CL8S is a fully symmetric version of CL8, obtained by adding to the latter

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 121–126.
springerlink.com

the duals (defined later) of the main rules: cocoupling, coweakening, and copulldown. The top-down symmetry of CL8S generates a number of interesting effects, some similar to those enjoyed by natural deduction systems. In this paper we show one such effect, i.e. deduction theorem. And we give a more strong result that the derivation (resp. proof) will be at most polynomially longer than the proof (resp. derivation).

Throughout the rest of this paper, unless otherwise specified, by a "rule" or a "derivation" etc. we mean one of the system CL8S. The reader is assumed to be familiar with the terminology, conventions and rules of CL8S that we will not reproduce.

2 Definition

Definition 1. *The* **dual** *of a conjunctive gate is a disjunctive gate, and vice verse.*

Definition 2. *The* **dual** *of a rule is obtained by interchanging premise with conclusion and conjunctive gates with disjunctive gates.*

Definition 3. *The* **dual** *of a derivation is obtained by turning it upside down and replacing each rule and each gate by its dual, respectively.*

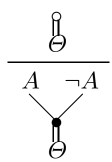
Definition 4. *The* **size** $|A|$ *of a cirquent* A *is the number of both ports and gates occurring in it.*

Definition 5. *The* **size** $|\triangle|$ *of a derivation* $\triangle$ *is the number of both ports and gates occurring in it.*

Definition 6. *A rule is* **strong admissible** *if, whenever a cirquent* B *follows from a cirquent* A *by that rule, there is also a derivation of* B *from* A.

3 Deduction Theorem

Extending the rule of (atomic) coupling from ports P, $\neg P$ to any cirquents A, $\neg A$, we obtain the following rule called **general coupling**.



As we will see, general coupling is strong admissible, which can be achieved by inductively replacing an application of it by applications on smaller cirquents.

Lemma 1. *The rule of general coupling is strong admissible, and reducing it to atomic coupling increases derivation sizes only polynomially.*

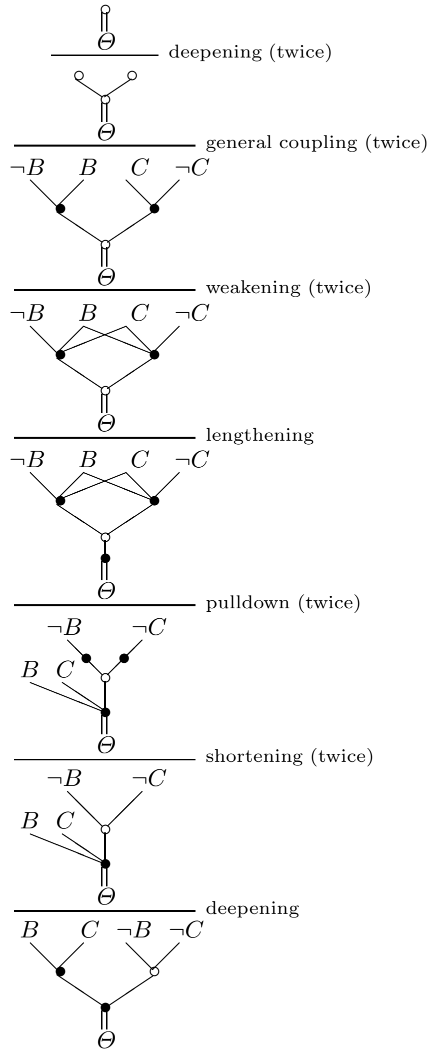
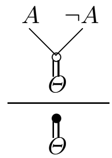


Fig. 1. A derivation that replacing an application of general coupling by applications on smaller cirquents

Proof. We make an induction on the structure of A. The cases when A is a childless gate or an atom are trivial: in the former case the derivation consists of one application of lengthening and one application of deepening, and in the latter case the application of general coupling is also an application of atomic coupling. We only have to consider the case when $A = B \vee C$, with the case when $A = B \wedge C$ being similar. We apply the induction hypothesis on the derivation, as shown in Figure 1. Obviously the length of the derivation is

$O(n)$, and so its size is $O(n^2)$, where $n = |A|$. Thus reducing general coupling to atomic one increases derivation sizes only polynomially. □

By taking the dual of general coupling, we obtain the following rule called **general cocoupling** whose strong admissibility is an immediate result from daulising the derivation of Lemma 1.

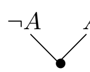


Lemma 2. *The rule of general cocoupling is strong admissible, and reducing it to atomic cocoupling increases derivation sizes only polynomially.*

We now see that one can easily move back and forth between a derivation and a proof of the corresponding implication via the deduction theorem:

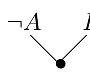
Theorem 1. (Deduction Theorem) *Let A, B be two cirquents. There is a derivation of B from A if and only if there is a proof of $A \to B$. And the derivation (resp. proof) will be at most polynomially longer than the proof (resp. derivation).*

Proof. ($\Rightarrow$:) The cirquent $A \to B$ is the abbreviation of $\neg A \vee B$. Assume that there is a derivation of B from A. Then a proof of $\neg A \vee B$ can be constructed by the following steps. Firstly, applying general coupling, we proceed from the axiom $\circ$ to the circuit



.

Secondly, duplicating the existing derivation of B from A, we continue proceeding from the cirquent above to the circuit



.

By Lemma 1, applying general coupling increases derivation sizes only polynomially. So such a proof of $\neg A \vee B$ is at most polynomially longer than the derivation of B from A.

($\Leftarrow$:) Assume that there is a proof of $\neg A \vee B$. we construct, top-down, a derivation of B from A, as shown in Figure 2. The vertical dots in Figure 2 stand for the duplication of the existing proof of $\neg A \vee B$. Obviously the length difference between this derivation and the proof of $\neg A \vee B$ is determined by the length of the derivation $\triangle$ which is the result of reducing general cocoupling to atomic version of this rule. By Lemma 2, the length of $\triangle$ is polynomial, and therefore, such a derivation of B from A is at most polynomially longer than the proof of $\neg A \vee B$. □

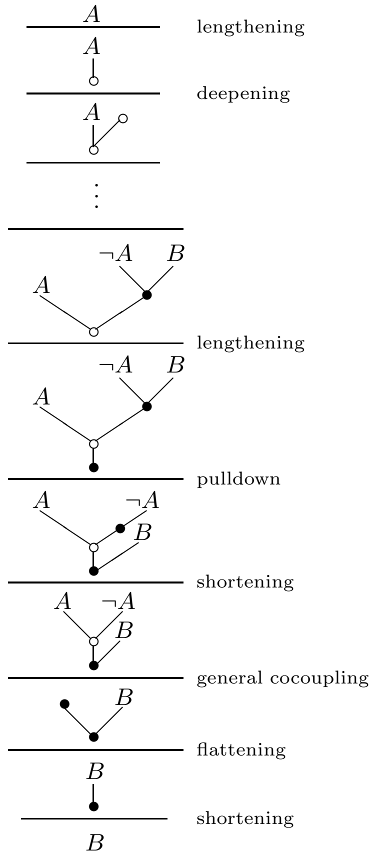


Fig. 2. A derivation of B from A

4 Conclusion

In this paper we prove the deduction theorem for the symmetric system CL8S of cirquent calculus, and give a more strong result that the derivation in the deduction theorem will be at most polynomially longer than the proof of implication, and vice versa.

Acknowledgements. This work described here is partially supported by the grants from the National Natural Science Foundation of China (NSFC No. 60974082).

References

1. Japaridze, G.: Cirquent calculus deepened. Journal of Logic and Computation 18, 983–1028 (2008)
2. Brünnler, K.: Deep Inference and Symmetry in Classical Proofs. Logos Verlag Press (2004)

3. Bruscoli, P., Guglielmi, A.: On the proof complexity of deep inference. ACM Transactions on Computational Logic 10, 1–34 (2009)
4. Japaridze, G.: Introduction to computability logic. Annals of Pure and Applied Logic 123, 1–99 (2003)
5. Japaridze, G.: Computability logic: a formal theory of interaction. In: Interactive Computation: The New Paradigm, pp. 183–223. Springer, Heidelberg (2006)
6. Japaridze, G.: In the beginning was game semantics. In: Games: Unifying Logic, Language, and Philosophy, pp. 249–350. Springer, Heidelberg (2009)
7. Gordeev, L., Haeusler, E.H., da Costa, V.G.: Proof compressions with circuit-structured substitutions. Journal of Mathematical Sciences 158, 645–658 (2009)
8. Guglielmi, A., Strassburger, L.: Non-commutativity and MELL in the calculus of structures. In: Fribourg, L. (ed.) CSL 2001 and EACSL 2001. LNCS, vol. 2142, pp. 54–68. Springer, Heidelberg (2001)
9. Guglielmi, A.: A system of interaction and structure. ACM Transactions on Computational Logic 8, 1–64 (2007)
10. Japaridze, G.: Introduction to cirquent calculus and abstract resorce semantics. Journal of Logic and Computation 16, 489–532 (2006)
11. Gentzen, G.: Investigations into logical deduction. American Philosophical Quartly 1, 288–306 (1964)

Equilateral Polygons in Classical Logic Metric Space

Ming-Di Hu

Institute of Mathematics, Shaanxi Normal University, Xi'an 710062, China
e-mail: humingdiwww@163.com
Department of Mathematics, Ankang College, Ankang 725000, China

Abstract. In the present paper, it is proved that some special graphs, such as equilateral polygons and right triangles, exist in the classical logic metric space. In addition, there are no equilateral triangles, length of whose lines is more than or equal to $\frac{2}{3}$, while the ones, length of whose lines can be arbitrarily close to $\frac{2}{3}$ or zero, do exist in the classical logic metric space. Lastly, it constructs an isometric reflexion transform, which can preserve the characters of the equilateral triangles, on the Lindenbaum algebra.

Keywords: Classical logic metric space, Equilateral triangle, Lindenbaum algebra, Reflexive transform.

1 Introduction

Mathematical logic, a subject of studying the form reasoning [1], offers approaches and methods of how to infer the conclusions from the known premises. This logic reasoning has been generally used in the artificial intelligence and related disciplines [2-5] and widely applied in many fields such as logic programming, automatic theorem proving, non-monotone logic and knowledge reasoning etc, which forms the theoretical basis for the modern computer science [6]. On the one hand, the characteristics of mathematical logic lie in formalization and symbolization, and focuses on form reasoning and strictly proving. On the other hand, the computational mathematics has the characteristics of flexibility and relaxation. In order to promote the application scope of mathematical logic, numerical calculation was introduced to the mathematical logic, by grading some basic concept, and the quantitative logic [2] was established by the second author of this paper. At the same time, the theory of logic metric space was established, and several approximate reasoning models were studied in this logic metric space. However, as to the structure of logic metric space, there are not many results yet, except

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 127–134.
springerlink.com

some logical topology characterizations given in [7] and topology characterizations of the sets of great harmony theories in propositional logic obtained in [8-10]. It can be said that the characteristics or properties of the logic metric space itself were still not very clear. In the reference [11], the structure of the logic metric space was discussed from a special angel, and it was proved that there exists a kind of reflexive transforms φ to be homomorphism mappings in the logic metric space, which can keep the logic equivalent relationship unchanged. Especially, φ^* , which can be naturally derived from φ , is a reflexive and isometric transform in the Lindenbaum algebra and a self-isomorphism. Furthermore, the characteristics and general form of the fixed points were obtained in the reference [11]. Based on the above, it is proved that in logic metric spaces there exist equilateral polygons and right triangles,etc. , and that the reflexive transform φ^* keeps equilateral polygons unchanged since φ^* is not only the self-isomorphism transform but also a isometric transform.

2 Basic Concepts

Definition 2.1. ([2]) *Let S be a countable set $p_1, p_2, p_3, \cdots$ of propositional variables, and $F(S)$ be the free algebra of type $(\neg, \vee, \rightarrow)$ generated by S, where $\neg$ is a unary operator, and $\vee$, $\rightarrow$ are binary operators on S. Members of S are called atoms and those of $F(S)$ are called well-formed formulas (wff's for short) or propositions.*

Suppose that $\{0,1\}$ is the simplest Boolean algebra, and for every $a, b \in \{0,1\}$

$$\neg a = 1 - a, a \vee b = max\{a, b\}, a \rightarrow b = 1 \quad iff \quad a \leq b \tag{1}$$

where $a \wedge b = \neg(\neg a \vee \neg b)$, then $\{0,1\}$ is also a free algebra of type $(\neg, \vee, \rightarrow)$.

Definition 2.2. ([2]) *Suppose that $A = A(p_1, p_2, \cdots, p_n)$ is a well-formed formula containing n atoms $p_1, p_2, \cdots, p_n$. Let p_i be replaced by $x_i, (i = 1, 2, \cdots, n)$, and the operators $\neg, \vee, \rightarrow$ followed from (1), then we obtain a Boolean function $f_A : \{0,1\}^n \rightarrow \{0,1\}$, which is called the Boolean function induced by A. And call $\frac{N(f)}{2^n}$ the truth degree of the formula A, written as $\tau(A)$, (where $N(f) = |f^{-1}(1)|$, please refer to [12]). In addition, define*

$$\xi(A, B) = \tau((A \rightarrow B) \wedge (B \rightarrow A)), A, B \in F(S), \tag{2}$$

and call $\xi(A, B)$ the similarity degree between A and B.

Besides, define

$$\rho(A, B) = 1 - \xi(A, B), A, B \in F(S), \tag{3}$$

then ρ is a pseudo-metric on $F(S)$ and call $(F(S), \rho)$ a logic metric space.

It is proved in [2] that $\xi(A, B) = 1$ (or $\rho(A, B) = 0$) if and only if $A \approx B$, (i.e., A and B are logically equivalent.).

Remark 2.1. *(i) The definition of the truth degree of formula is not its original definition in the reference* [2]*, but they are equivalent according to the reference* [12]*.*

(ii) The indexes of atoms in formula A are unnecessarily from 1 to n continually, but we suppose that the maximal index is n, and let $B = A \vee (p_1 \wedge \neg p_1 \vee \cdots \vee (p_n \wedge \neg p_n)$. Then B and A are logically equivalent, meanwhile the indexes of atoms in formula B continue from 1 to n, and the Boolean function induced by A can also be written as $f(x_1, \cdots, x_n)$. So, we always suppose that the indexes of atoms in formula A are continuous from 1 to n in the following context when the maximal index of all atoms of the formula A is n.

(iii) $\rho(A, B) = 0$ when A and B is similar (i.e., $\xi(A, B) = 1$), but A and B may be different formulas at the mean time. From this, $(F(S), \rho)$ is not a metric space and only is a pseudo metric space. However, ρ induces a real metric ρ^*, also written as ρ, in Lindenbaum quotient algebra $[F(S)] = F(S)/\approx$ in a natural way, and $([F(S), \rho])$ is a metric space,so we call $([F(S), \rho])$ as the classical logical metric space.

Definition 2.3. ([11]) *Suppose that $S = \{P_1, P_2, \cdots\}$ is the set of all atomic formulas, and $F(S)$ is the set of all Classical propositional logic formulas. Define $\varphi : F(S) \to F(S)$ as following:*

Suppose $A = f(P_1, P_2, \cdots p_n) \in F(S)$, let $\varphi(A) = f(\neg P_1, \neg P_2, \cdots, \neg p_n)$. We call this transformation a reflection transformation.

Theorem 2.1. ([11]) *$\varphi : F(S) \to F(S)$ is a homomorphism transform.*

Theorem 2.2. ([11]) *Assume that A is logically equivalent with B, then $\varphi(A)$ is logically equivalent with $\varphi(B)$ too.*

Definition 2.4. ([8]) *Suppose that $(F(S), \rho)$ is a logical metric space, let*

$$\rho^*([A], [B]) = \rho(A, B), A, B \in F(S).$$

Then $([F(S)], \rho^*)$ is a real metric space,also written as $([F(S)], \rho)$ briefly.

Definition 2.5. ([11]) *Suppose that $\varphi : F(S) \to F(S)$ is a reflexive transform on $F(S)$,let*

$$\varphi^*([A]) = [\varphi(A)], A \in F(S). \tag{4}$$

We call φ^* as a reflexive transform on $[F(S)]$ too.

Theorem 2.3. ([11]) *$\varphi^* : [F(S)] \to [F(S)]$ is an automorphism transformation in the Lindenbaum algebras.*

Definition 2.6. ([11]) *Suppose that (M, d) is a metric space, $f : M \to M$ is a transform on M. If*

$$d(f(x), f(y)) = d(x, y), x, y \in M, \tag{5}$$

then we call f as a isometric transform on M.

Theorem 2.4. ([11]) *$\varphi^* : ([F(S)], \rho) \to ([F(S)], \rho)$ is a isometric transform.*

3 Equilateral Polygons on the Classical Logic Metric Space

In this section, the existence of equilateral polygons were discussed. From the issue, we obtain the conclusions that the reflexive transform φ^* keep the especial relationship of the lines and the angles unchanged. Hereinafter, though we discussed all issues in the classical logic metric space $([F(S)], \rho)$, there are no differences of the relationships of measure between the classical logic metric space $([F(S)], \rho)$ and the logical metric space $(F(S), \rho)$. To facilitate the writing, we note the equivalence classes $[A]$ as A in the text below.

Theorem 3.1. *There must be some equilateral triangles on the classical logical metric space $([F(S)], \rho)$, but there are no any equilateral triangles, the length of whose lines is longer than $\frac{2}{3}$.*

Proof. In the logical metric space $([F(S)], \rho)$, we let $A = p, B = q, C = r$. According to the Equations (2), (3), we have

$$\rho(A, B) = \rho(A, C) = \rho(C, B) = \frac{1}{2}. \tag{6}$$

That is to say the triangle ABC is an equilateral triangle, the length of whose lines is equal to $\frac{1}{2}$. Therefore there must be some equilateral triangles on the logical metric space $([F(S)], \rho)$.

Secondly, suppose that there exists an equilateral triangle, the length of whose lines is longer than $\frac{2}{3}$, namely, $\rho(A, B) > \frac{2}{3}, \rho(A, C) > \frac{2}{3}, \rho(C, B) > \frac{2}{3}$, then

$$\xi(A, B) < \frac{1}{3}, \xi(A, C) < \frac{1}{3}, \xi(C, B) < \frac{1}{3}. \tag{7}$$

Such as the note (2) above, it may be assumed that A, B, and C contain same n atomics such as $p_1, p_2, \cdots, p_n$. They induce three Boolean functions with n variables, and the three functions are noted as f_A and f_B or f_C, respectively. Now, we have the fact that the similarity degree between arbitrary two formulas of A, B, and C is less than $\frac{1}{3}$ while the sequence pairs $(x_1, x_2, \cdots, x_n)$ is given any value in $\{0, 1\}^n$. In another words, there are less than $\frac{1}{3}$ values that make the equality $f_A = f_B$ hold, and so do the equalities $f_A = f_C$ and $f_C = f_B$. Therefore, there must be at least one value v in $\{0, 1\}^n$ making the following three inequalities hold at the same time:

$$f_A \neq f_B, f_A \neq f_C, f_C \neq f_B. \tag{8}$$

In fact, f_A, f_B and f_C can be zero or one, so it is impossible that the three inequalities hold at the same time. Contradictions!

Theorem 3.2. *There is not any equilateral triangle, the length of whose lines is equal to $\frac{2}{3}$ on the logical metric space $([F(S)], \rho)$.*

Proof. Suppose that there exists an equilateral triangle $\triangle ABC$, the length of whose lines is equal to $\frac{2}{3}$, namely,$\rho(A,B) = \rho(A,C) = \rho(C,B) = \frac{2}{3}$,then $\xi(A,B) = \frac{1}{3}$.

According to the equality (2), we have

$$\xi(A,B) = \tau((A \to B) \wedge (B \to A)) = \frac{1}{3}. \tag{9}$$

The equality is impossible to hold, because the set of truth degree of all formulas in $F(S)$ is as following $H = \{\frac{k}{2^n} | k = 0, 1, \cdots, 2^n; n = 1, 2, \}$. (Please refer to the proposition 9.1.1 in the reference [2]). If the equality (9) holds, then there exists an integer k such that the equality $\frac{k}{3k} = \frac{k}{2^n}$ holds, namely, $k = \frac{2^n}{3}$. However, this integer k only equals every integer from 0 to 2^n. Contradictions.

Theorem 3.3. *Suppose that ε is a positive number arbitrary given, then there exist some equilateral triangles, the length of whose lines is less than ε on the classical logical metric space $([F(S)], \rho)$.*

Proof. let

$$A = p_1 \wedge p_2 \wedge \cdots \wedge p_n, B = q_1 \wedge q_2 \wedge \cdots \wedge q_n, C = r_1 \wedge r_2 \wedge \cdots \wedge r_n.$$

Then, from the equality (2) we have

$$\xi(A,B) = \tau((A \to B) \wedge (B \to A)) = \frac{(2^n - 1)^2 + 1}{2^{2n}}. \tag{10}$$

It is easy to prove that $\frac{(2^n-1)^2+1}{2^{2n}} \to 1$ when $n \to \infty$, namely, for any given positive number ε, there exists a positive integer N such that when $n > N$,

$$\rho(A,B) < \varepsilon.$$

Similarly, we also have

$$\rho(A,C) = \frac{(2^n-)^2+1}{2^{2n}} < \varepsilon, \rho(A,C) = \frac{(2^n-)^2+1}{2^{2n}} < \varepsilon.$$

That is to say, there exist some equilateral triangles the length of whose lines is less to ε on the classical logical metric space.

Corollary 3.1. *There must be some equilateral polygons on the classical logical metric space $([F(S)], \rho)$, and the length of their lines is arbitrarily close to zero.*

Proof. It is proved by the way which is similar to Theorem 3.1 and Theorem 3.3.

Theorem 3.4. *There exists a right triangle on the classical logical metric space $([F(S)], \rho)$.*

Proof. Let

$$A = p_1 \wedge p_2, B = p_1 \vee p_2, C = A = p_3 \wedge p_4.$$

According to the equalities (2), (3) we have

$$\xi(A,B) = \tfrac{4}{8}, \xi(A,C) = \tfrac{5}{8}, \xi(C,B) = \tfrac{3}{8};$$

then

$$\rho(A,B) = \tfrac{4}{8}, \rho(A,C) = \tfrac{3}{8}, \rho(C,B) = \tfrac{5}{8}.$$

It is clear to see that $\triangle ABC$ is a right triangle since it satisfies the Pythagorean theorem.

Lemma 3.1. *For any natural number k, $2^{2k}+2$ can be divisible by 3, and the following equality holds.*

$$2^{2k} = 2 \times \frac{2^{2k}+2}{3} + (\frac{2^{2k}+2}{3} - 2). \tag{11}$$

Proof. It is proved using mathematical induction. The equality $2^{2n}+2 = 3\times 2$ holds when n equals 1, so this proposition holds. Now suppose that this proposition holds when n is $k-1$ (k is arbitrary natural number), namely, there exists a positive integer h_1 such that the equality $2^{2(k-1)}+2 = 3\times h_1$ holds. Therefore, from the equality as following:

$$2^{2k}+2 = 2^{2(k-1)} \times 2^2 + 2 = (3h_1 - 2) \times 4 + 2 = 12h_1 - 6 = 3 \times (4h_1 - 2).$$

We have that this proposition holds for k. It is clear to see the proposition holds for any natural number k, namely,$2^{2k}+2$ can be divisible by 3. And, it is easy to prove that the equality (11) holds.

Theorem 3.5. *For any positive number ε which can be arbitrarily close to zero, there exists an equilateral triangle on the classical logical metric space $([F(S)], \rho)$, the length of whose lines is between $\frac{2}{3}-\varepsilon$ and $\frac{2}{3}$.*

Proof. For an arbitrary $\varepsilon > 0$, let k be large enough so that $\frac{2}{3\times 2^{2k}} < \varepsilon$. Now we argue the Boolean function $f_A(x_1, x_2, \cdots, x_n)$ and $f_B(x_1, x_2, \cdots, x_n)$ or $f_C(x_1, x_2, \cdots, x_n)$ with $n = 2k$ variables, where $(x_1, x_2, \cdots, x_{2k}) \in \{0,1\}^{2k}$. We make the self-variable of vectors arranged according to lexicographic sequence. Firstly, let all vectors equal zero, namely, let $f_A(x_1, x_2, \cdots, x_{2k}) \equiv 0$, so we have a formula $A(p_1, p_2, \cdots, p_{2k})$ which is $\bar{0}$. Secondly, we change a part of values of the Boolean function above, that is to say, let these vectors with lexicographic sequence equal to zero from the first to the $\frac{2^{2k}+2}{3}$'th, the others equal to one, because the number $\frac{2^{2k}+2}{3}$ is positive integer according to Lemma 3.6. So, we obtain the second Boolean function $f_B(x_1, x_2, \cdots, x_{2k})$ and the formula $B(p_1, p_2, \cdots, p_{2k})$ with $2k$ atoms which is corresponding to the second Boolean. Thirdly, let these sequent vectors equal zero from $\frac{2^{2k}+2}{3}$'th to $2\times\frac{2^{2k}+2}{3}+1$'th, and the others equal 1, so we obtain also the third Boolean function $f_C(x_1, x_2, \cdots, x_{2k})$ and the third formula $C(p_1, p_2, \cdots, p_{2k})$ with $2k$ atoms which is corresponding to the third Boolean function. By the equality (2)and (3) , as well as the construction of these above Boolean functions, it is not difficult to test as following:

$\frac{1}{3} < \xi(A,B) = \xi(B,C) = \xi(C,A) = \frac{1}{3} + \frac{2}{3+2^{2k}} < \frac{1}{3} + \varepsilon,$
namely,
$\frac{2}{3} > \rho(A,B) = \rho(B,C) = \rho(C,A) = \frac{2}{3} - \frac{2}{3+2^{2k}} > \frac{2}{3} - \varepsilon.$ So this proposition is proved.

Corollary 3.2. *For any positive number ε which can be arbitrarily close to zero, there exist numerous equilateral triangles on the classical logical metric space $([F(S)], \rho)$, the length of whose lines is between $\frac{2}{3} - \varepsilon$ and $\frac{2}{3}$.*

Proof. We can see from the proof of Theorem 3.7 that there exist numerous k which satisfies the condition $\frac{2}{3+2^{2k}} < \varepsilon$, therefore, we can draw many equilateral triangles like above according to the way of the proof of Theorem 3.7.

4 The Characters of the Equidistant Transform on the Space of $([F(S)], \rho)$

Theorem 4.1. *The isometric transform φ^* preserves the shape of equilateral triangles on the classical logical metric space $([F(S)], \rho)$.*

Proof. Suppose that $\triangle ABC$ is an equilateral triangle, from the definition of reflective transform we have

$$\varphi^*(A) = A', \varphi^*(B) = B', \varphi^*(C) = C',$$

then the triangle $\triangle A'B'C'$ have the following character according to Theorem 2.10,

$$\rho(A', B') = \rho(\varphi^*(A), \varphi^*(B)) = \rho(A, B).$$

In the same way, we have $\rho(A', C') = \rho(A, C)$, $\rho(B', C') = \rho(B, C)$, therefore, so is $\triangle A'B'C'$. Because a triangle can decide a plane and equal lines make relative angles equal, the isometric transform preserves the shape of every equilateral triangle on $([F(S)], \rho)$.

It is easy to obtain the following corollary by Theorem 2.10 and Corollary 3.4.

Corollary 4.1. *The isometric transform preserves the length of lines of every equilateral polygon on $([F(S)], \rho)$, but the shape of any equilateral polygon is not necessarily preserved by this isometric transform.*

Theorem 4.2. *The isometric transform can preserve the shape of every right triangle on the metric space $([F(S)], \rho)$.*

Proof. It can be proved similarly to Theorem 4.1, only using the character of the isometric transform φ^* .

5 Conclusion

We discussed the structure of some logic metric spaces from a special angle in the reference [11], and it is proved that there is a reflexive transform φ in the classical logic metric space, which is a homomorphic mapping and makes the logic of equivalence relation unchanged. And, φ^* is a reflexive transform of Lindenbaum algebra derived from φ, which is an automorphic and isometric transform of Lindenbaum algebra. Besides, the general form of fixed points can be obtained by studying the features of fixed points. In this paper, on the basis of above paper, it is further proved that there exist some special graphics such as equilateral triangles and polygons et. on the logic metric space. And it is also proved that an isometric φ^* in Lindenbaum algebras can preserve the shape of every equilateral triangle and the length of lines of every equilateral polygon. We will discuss the other issues of reflection transform furthermore in the next paper, for example, how to discuss and seek the structure of classical logic metric space, and how to discuss the result in the Lukasiewicz system or Ro system, etc.

Acknowledgements. This work described here is partially supported by the grant from the National Science Foundation of China (NSFC No. 10771129), and Innovation Funds of Graduate Programs, SNU.(IFGPS No. 2009CXB006).

References

1. Hamilton, A.G.: Logic for Mathematicians. Cambridge University Press, London (1978)
2. Wang, G.J.: Introduction to Mathematical Logic and Resolution Principle, 2nd edn. Science in China Press, Beijing (2006)
3. Gottwald, S.: A treatise on many-valued logics. Research Studies Press, Baldock (2001)
4. Pei, D.W.: A survey of ten years' studies on fuzzy logic and fuzzy reasoning. Journal of Engineering Mathematics 21(2), 249–258 (2004)
5. Wang, Y.Y.: Logic of computer science. Science in China Press, Beijing (1989)
6. Cao, L.Z., Wei, C.F.: Modern computer's application and improvement. Science in China Press, Beijing (2002)
7. Wang, G.J., She, Y.H.: Topological description of divergency and consistency of two-valued propositional theories. Acta Mathematica Sinica 50(4), 1–10 (2007)
8. Wang, G.J., Wang, W.: Logical metric spaces. Acta Mathematica Sinica 44(1), 159–168 (2000)
9. Wang, G.J., Wang, W., Song, J.S.: Topology on the set of maximal consistent propositional theories and the Cantor ternary set. Journal of Shaanxi Normal University: Natural Science Edition 35(2), 1–5 (2007)
10. Zhou, H.J., Wang, G.J.: Characterizations of maximal consistent theories in the formal deductive system $\mathcal{L}^*$ (NM-logic) and Cantor Space. Fuzzy Sets and Systems 158(23), 2591–2604 (2007)
11. Hu, M.D., Wang, G.: Reflexive transform in classical logic metric space. Journal of Shaanxi Normal University: Natural Science Edition 37(6), 1–4 (2009)
12. Wen, Q.Y., Niu, X.X., Yang, Y.X.: Boolean Functions in Current Cryptology. Science in China Press, Beijing (2000)

Reflexive Transformation in L^*-Logic Metric Space

Qing-Ping Wang[1,2] and Guo-Jun Wang[1]

[1] College of Mathematics and Information Science, Shaanxi Normal University, Xi'an 710062, China
e-mail: lc_wqp@163.com, gjwang@snnu.edu.cn

[2] School of Mathematical Sciences, Liaocheng University, Liaocheng 252059, China

Abstract. In this paper, a reflexive transformation φ on $F(S)$ in the logic system L^* is proposed, which is an automorphism of $F(S)$. It is proved that the concepts of truth degree, similarity degree and pseudo-distance are remain under the reflexive transformation φ. Moreover, three different approximate reasoning patterns are introduced in the system L^*, and it is also proved that for any $A \in F(S)$ and $\Gamma \subset F(S)$, A is an approximate conclusion of Γ if and only if $\varphi(A)$ is an approximate conclusion of $\varphi(\Gamma)$ in the sense of any one of the three patterns. Finally, the properties of reflexive transformation are investigated and relation of fixed points in classical logic system L and those in L^* are studied.

Keywords: Logic system L^*, reflexive transformation, pseudo-distance, approximate reasoning, fixed point.

1 Introduction

Mathematical logic is a subject dealing with formalized reasoning, in which the methods on inferring the desirable conclusion from the known premise are studied. This kind of logic reasoning has been widely applied in artificial intelligence and related topics [1-5]. The logic system L^* is proposed by the second author of this paper, in which the completeness theorem holds, i.e., the syntax and the semantics are in perfect harmony. Additionally, the concepts of similarity degree and pseudo-distance are also given in L^*, and several kinds of approximate reasoning are proposed. Until now, a relatively complete and mature quantitative logic has been formed [1]. The symbolization and formalization of mathematical logic and the numerical computation of computational mathematics are connected in quantitative logic. It enables mathematical logic to have some kind of flexibility and thus extends the scope

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 135–144.
springerlink.com

of possible applications [1,6-14]. In paper [15], the reflexive transformation is introduced in classical logic metric space, which enriches the theory of quantitative logic. In this paper, we aim to extend this result to the environment of fuzzy logic system L^* and a reflexive transformation φ is introduced in the logic system L^*. It has been proved that the concepts of truth degree, similarity degree and pseudo-distance enjoy the following properties under the transformation φ:

$\tau_{R_\infty}(A) = \tau_{R_\infty}(\varphi(A)), \xi_R(A,B) = \xi_R(\varphi(A),\varphi(B)), \rho_R(A,B) = \rho_R(\varphi(A),\varphi(B))$.

Moreover,three different approximate reasoning patterns are introduced in system L^*. And it is proved that for any $A \in F(S)$ and $\Gamma \subset F(S)$, A is an approximate conclusion of Γ if and only if $\varphi(A)$ is that of $\varphi(\Gamma)$ under the proposed reasoning patterns. Finally, the properties of the reflexive transformation φ_2^* on $[F]$ in logic system L^* are investigated and the relationship of fixed points in L and those in L^* is studied.

2 Preliminaries

Definition 1. ([1]) *Let $S = \{p_1, p_2, \cdots\}$, and $F(S)$ be a free algebra of type $(\neg, \vee, \rightarrow)$ generated by S. Each member of $F(S)$ is called a formula (also called proposition) of L^*, and each member of S is called an atomic formula (atomic proposition) of L^*.*

Definition 2. ([1]) *Let $\upsilon : F(S) \rightarrow [0,1]$ be a mapping from the set $F(S)$ of all formulas of L^* into the R_0-unit interval $[0,1]$. υ is called a valuation of $F(S)$ in the R_0-unit interval $[0,1]$(briefly, valuation) if υ is a homomorphism of type $(\neg, \vee, \rightarrow)$, i.e. $\upsilon(\neg A) = \neg\upsilon(A)$, $\upsilon(A \vee B) = \upsilon(A) \vee \upsilon(B) = max\{\upsilon(A), \upsilon(B)\}$, $\upsilon(A \rightarrow B) = \upsilon(A) \rightarrow \upsilon(B) = R_0(\upsilon(A), \upsilon(B))$.*

The set of all valuations of $F(S)$ will be denoted by $\bar{\Omega}$.

Definition 3. ([1]) *Let $A, B \in F(S)$.*
(i)A is called a tautology, denoted by $\models A$, if $\upsilon(A) = 1$ for every $\upsilon \in \bar{\Omega}$. A is called a contradiction , if $\upsilon(A) = 0$ for every $\upsilon \in \bar{\Omega}$.
(ii)A and B are said to be logically equivalent, in symbols, $A \approx B$, if $\upsilon(A) = \upsilon(B)$ for every $\upsilon \in \bar{\Omega}$.

Definition 4. ([1]) *Let $A(p_1, \cdots, p_n)$ be a formula containing n atomic propositions $p_1, \cdots, p_n$ by using the logical connectives $\neg, \vee$ and $\rightarrow$, and $(x_1, \cdots, x_n) \in [0,1]^n$. Substitute x_i for p_i in $A(i = 1, \cdots, n)$, and keep the logical connectives in A unchanged but explaining them as the corresponding operations defined by $\neg x_1 = 1 - x_1, x_1 \vee x_2 = max\{x_1, x_2\}, x_1 \rightarrow x_2 = R_0(x_1, x_2)$. Then we get a $n-$ray function $\bar{A}(x_1, \cdots, x_n)$, called the R_0 function induced by A.*

Definition 5. ([1]) *Let $A, B \in F(S)$. A and B are said to be provably equivalent, denoted by $A \sim B$, if both $A \rightarrow B$ and $B \rightarrow A$ are theorems.*

Theorem 1. ([1])*(Completeness theorem of L^*) Let $A \in F(S)$. Then A is a theorem in L^* if and only if A is a tautology, i.e. $\vdash A$ if and only if $\models A$.*

3 Reflexive Transformation

Definition 6. ([1]) *Let $S = \{p_1, p_2, \cdots\}$ be a set of all atomic formulas, and $F(S)$ be a set of all formulas of L^*. A mapping $\varphi : F(S) \to F(S)$ is called a reflexive transformation of $F(S)$, if $\forall A \in F(S)$, let $A = A(p_1, \cdots, p_n)$, then $\varphi(A) = A(\neg p_1, \cdots, \neg p_n)$.*

Theorem 2. *The reflexive transformation $\varphi : F(S) \to F(S)$ is an automorphism of $F(S)$.*

Proof. Let $A = A(p_1, \cdots, p_n) \in F(S), B = B(q_1, \cdots, q_m) \in F(S)$, then $\varphi(\neg A) = \varphi(\neg A(p_1, \cdots, p_n)) = \neg A(\neg p_1, \cdots, \neg p_n) = \neg\varphi(A)$, $\varphi(A \vee B) = \varphi(A(p_1, \cdots, p_n) \vee B(q_1, \cdots, q_m)) = A(\neg p_1, \cdots, \neg p_n) \vee B(\neg q_1, \cdots, \neg q_m) = \varphi(A) \vee \varphi(B)$, $\varphi(A \to B) = \varphi(A(p_1, \cdots, p_n) \to B(q_1, \cdots, q_m)) = A(\neg p_1, \cdots, \neg p_n) \to B(\neg q_1, \cdots, \neg q_m) = \varphi(A) \to \varphi(B)$.

Hence φ is a homomorphism of $F(S)$.

$\forall C, D \in F(S)$, if $\varphi(C) = \varphi(D) = E(p_1, \cdots, p_t)$, then $C = D = E(\neg p_1, \cdots, \neg p_t)$. Hence φ is an injection.

$\forall E(p_1, \cdots, p_t) \in F(S)$, there exists $E(\neg p_1, \cdots, \neg p_t) \in F(S)$, such that $\varphi(E(\neg p_1,$
$\cdots, \neg p_t)) = E$. Hence φ is a surjection.

Therefore the reflexive transformation φ is an automorphism of $F(S)$.

Theorem 3. *If A and B are logically equivalent, then $\varphi(A)$ and $\varphi(B)$ are logically equivalent.*

Proof. Assume that A and B are logically equivalent. Let $A = A(p_1, \cdots, p_n), B = B(p_1, \cdots, p_n).\bar{A}, \bar{B}$ are R_0 functions induced by A and B respectively. For any valuation $\upsilon \in [0,1]^n$, there exists a valuation $\mu = 1-\upsilon \in [0,1]^n$ such that $\bar{A}(\mu(p_1), \cdots, \mu(p_n)) = \mu(A) = \mu(B) = \bar{B}(\mu(p_1), \cdots, \mu(p_n))$, Moreover, since $\varphi(A) = A(\neg p_1, \cdots, \neg p_n), \varphi(B) = B(\neg p_1, \cdots, \neg p_n)$, then $\upsilon(\varphi(A)) = \bar{A}(\upsilon(\neg p_1), \cdots, \upsilon(\neg p_n)) = \bar{A}(1 - \upsilon(p_1), \cdots, 1 - \upsilon(p_n)) = \bar{A}(\mu(p_1), \cdots, \mu(p_n)) = \bar{B}(\mu(p_1), \cdots, \mu(p_n)) = \bar{B}(1 - \upsilon(p_1), \cdots, 1 - \upsilon(p_n)) = \bar{B}(\upsilon(\neg p_1), \cdots, \upsilon(\neg p_n)) = \upsilon(\varphi(B))$.

And hence $\varphi(A)$ and $\varphi(B)$ are logically equivalent due to the arbitrariness of υ.

Corollary 1. *If A and B are provably equivalent, then $\varphi(A)$ and $\varphi(B)$ are also provably equivalent.*

Theorem 4. *Let $A \in F(S)$.*
(i)A is a tautology if and only if $\varphi(A)$ is a tautology.
(ii)A is a contradiction if and only if $\varphi(A)$ is a contradiction.

Proof. (i) Assume that $A = A(p_1, \cdots, p_n)$.Since A is a tautology, for any valuation $\upsilon \in [0,1]^n$, there exists a valuation $\mu = 1 - \upsilon \in [0,1]^n$ such that $\mu(A) = \bar{A}(\mu(p_1), \cdots, \mu(p_n)) = 1$. Since $\varphi(A) = A(\neg p_1, \cdots, \neg p_n)$, then

$v(\varphi(A)) = \bar{A}(1-v(p_1),\cdots,1-v(p_n)) = \bar{A}(\mu(p_1),\cdots,\mu(p_n)) = 1$. And therefore $\varphi(A)$is a tautology due to the arbitrariness of v.

Notice that $\varphi(\varphi(A)) = A$. Hence, if $\varphi(A)$ is a tautology, then A is a tautology.

The proof of (ii) is similar to (i), and hence is omitted here.

Corollary 2. *Let $A \in F(S)$, A is a theorem if and only if $\varphi(A)$ is a theorem.*

4 Truth Degree, Similarity Degree and Ppseudo-Distance of L^* under the Reflexive Transformation φ

Definition 7. ([1]) *Let $A = A(p_1,\cdots,p_n)$ be a formula containing n atomic formulas $p_1,\cdots,p_n$ in L^*. Define*

$$\tau_{R_\infty}(A) = \int_0^1 \cdots \int_0^1 \bar{A}_\infty(x_1,\cdots,x_n)dx_1\cdots dx_n, \tag{1}$$

$\tau_{R_\infty}(A)$is called the integrated truth degree of A in L^*, where $\bar{A}_\infty : [0,1]^n \to [0,1]$ is a function induced by A in the usual way.

Proposition 1. *$\forall A \in F(S)$, let φ be the reflexive transformation of $F(S)$ as defined in Definition 6, then*

$$\tau_{R_\infty}(A) = \tau_{R_\infty}(\varphi(A)). \tag{2}$$

Proof. $\forall A \in F(S)$, assume that $A = A(p_1,\cdots,p_n)$, then $\tau_{R_\infty}(A) = \int_0^1 \cdots \int_0^1 \bar{A}_\infty(x_1,\cdots,x_n)dx_1\cdots dx_n$.

Since $\varphi(A) = A(\neg p_1,\cdots,\neg p_n)$, then $\tau_{R_\infty}(\varphi(A)) = \int_0^1 \cdots \int_0^1 \bar{A}_\infty(1-x_1,\cdots,1-x_n)dx_1\cdots dx_n = (-1)^n \int_0^1 \cdots \int_0^1 \bar{A}_\infty(1-x_1,\cdots,1-x_n)d(1-x_1)\cdots(1-dx_n) = (-1)^n \int_1^0 \cdots \int_1^0 \bar{A}_\infty(y_1,\cdots,y_n)dy_1\cdots dy_n = \int_0^1 \cdots \int_0^1 \bar{A}_\infty(y_1,\cdots,y_n)dy_1\cdots dy_n = \tau_{R_\infty}(A)$.

Definition 8. ([1]) *In L^*, let $A, B \in F(S)$. Define*

$$\xi_R(A,B) = \tau_{R_\infty}((A \to B) \wedge (B \to A)), \tag{3}$$

$\xi_R(A,B)$ is called the similarity degree between A and B.

Proposition 2. *In L^*, $\forall A, B \in F(S)$, let φ be the reflexive transformation of $F(S)$ as defined in Definition 6. Then*

$$\xi_R(A,B) = \xi_R(\varphi(A),\varphi(B)) \tag{4}$$

Proof. We have from Proposition 1 that,
$\tau_{R_\infty}((A \to B) \wedge (B \to A)) = \tau_{R_\infty}(\varphi((A \to B) \wedge (B \to A)))$,
i.e.,
$\tau_{R_\infty}((A \to B) \wedge (B \to A)) = \tau_{R_\infty}((\varphi(A) \to \varphi(B)) \wedge (\varphi(B) \to \varphi(A)))$.
Hence $\xi_R(A, B) = \xi_R(\varphi(A), \varphi(B))$.

Definition 9. ([1]) *In L^*, a function $\rho_R : F(S) \times F(S) \to [0, 1]$ is defined as following:*

$$\rho_R(A, B) = 1 - \xi_R(A, B), \quad A, B \in F(S). \tag{5}$$

Then ρ_R is a pseudo-metric on $F(S)$, called the natural pseudo-metric (sometimes simply called pseudo-metric) on $F(S)$ in L^*.

Proposition 3. *In L^*, $\forall A, B \in F(S)$, let φ be the reflexive transformation of $F(S)$ as defined in Definition 6, then*

$$\rho_R(A, B) = \rho_R(\varphi(A), \varphi(B)). \tag{6}$$

The proof of Proposition 3 follows immediately from Proposition 2 and Definition 9, and hence is omitted here.

5 Approximate Reasoning of L^* under the Reflexive Transformation φ

Definition 10. *$\Gamma \subset F(S)$, let φ be the reflexive transformation of $F(S)$ as defined in Definition 6. Define*

$$\varphi(\Gamma) = \{\varphi(A) | A \in \Gamma\}. \tag{7}$$

Theorem 5. *$\Gamma \subset F(S)$, let φ be the reflexive transformation of $F(S)$ as defined in Definition 6. Then $A \in D(\Gamma)$ if and only if $\varphi(A) \in D(\varphi(\Gamma))$.*

Proof. $\forall A \in D(\Gamma)$, i.e. there exists a finite sequence of formulas $A_1, \cdots, A_n = A$ such that for each i with $1 \le i \le n$, either A_i is a theorem of L^*, or $A_i \in \Gamma$, or there exist $j, k < i$ such that A_i follows from A_j and A_k by MP.

If A_i is a theorem of L^*, then we have from Corollary 1 that $\varphi(A)$ is theorem.

If $A_i \in \Gamma$, then $\varphi(A_i) \in \varphi(\Gamma)$.

If there exist $j, k < i$ such that A_i follows from A_j and A_k by MP. Without loss of generality, we assume that $A_k = A_j \to A_i$, then $\varphi(A_j), \varphi(A_k) = \varphi(A_j \to A_i) = \varphi(A_j) \to \varphi(A_i) \in \varphi(\Gamma)$, and $\varphi(A_i)$can be obtained by MP. i.e. $\varphi(A_i) \in D(\varphi(\Gamma))$. Therefore $\varphi(A) \in D(\varphi(\Gamma))$.

For the converse direction, it can be proved immediately by using $\varphi(\varphi(A)) = A$.

Remark 1. From Theorem 5, if A is a conclusion of Γ, then $\varphi(A)$ is a conclusion of $\varphi(\Gamma)$. In converse, if $\varphi(A)$ is a conclusion of $\varphi(\Gamma)$, then A is a conclusion of Γ.

Example 1. Let $\Gamma = \{p_2, \neg p_1 \to \neg p_2\}$, then $\neg p_1 \in D(\varphi(\Gamma))$.

Proof. (1)$\neg p_1 \to \neg p_2$ (Γ)
(2)$(\neg p_1 \to \neg p_2) \to (p_2 \to p_1)$ (L^*2)
(3)$p_2 \to p_1$ (1)(2)MP
(4)p_2 (Γ)
(5)p_1 (1)(2)MP
Hence $p_1 \in D(\Gamma)$. From Theorem 5, we have $\neg p_1 \in D(\varphi(\Gamma))$.

Definition 11. ([1]) *In L^*, let $\Gamma \subset F(S), B \in F(S), \varepsilon > 0$. If*

$$\rho_R(B, D(\Gamma)) < \varepsilon, \tag{8}$$

then B is called a conclusion of Γ of I-type with error less than ε, denoted by $B \in D^1_{L^*,\varepsilon}(\Gamma)$.

Proposition 4. *In L^*, assume that $\Gamma \subset F(S), B \in F(S), \varepsilon > 0$. Let φ be the reflexive transformation of $F(S)$ as defined in Definition 6. Then $B \in D^1_{L^*,\varepsilon}(\Gamma)$ if and only if $\varphi(B) \in D^1_{L^*,\varepsilon}(\varphi(\Gamma))$.*

Proof. $B \in D^1_{L^*,\varepsilon}(\Gamma) \Longleftrightarrow \rho_R(B, D(\Gamma)) < \varepsilon \Longleftrightarrow$ There exists $A \in D(\Gamma)$ such that $\rho_R(A, B) < \varepsilon \Longleftrightarrow$ There exists $\varphi(A) \in D(\varphi(\Gamma))$ such that $\rho_R(\varphi(A), \varphi(B)) < \varepsilon \Longleftrightarrow \varphi(B) \in D^1_{L^*,\varepsilon}(\varphi(\Gamma))$.

Definition 12. ([1]) *In L^*, let $\Gamma \subset F(S), B \in F(S), \varepsilon > 0$. If*

$$1 - sup\{\tau_{R_\infty}(A \to B) | A \in D(\Gamma)\} < \varepsilon, \tag{9}$$

then B is called a conclusion of Γ of II-type with error less than ε, denoted by $B \in D^2_{L^*,\varepsilon}(\Gamma)$.

Proposition 5. *In L^*, assume that $\Gamma \subset F(S), B \in F(S), \varepsilon > 0$. Let φ be the reflexive transformation of $F(S)$ as defined in Definition 6. Then $B \in D^2_{L^*,\varepsilon}(\Gamma)$ if and only if $\varphi(B) \in D^2_{L^*,\varepsilon}(\varphi(\Gamma))$.*

Definition 13. ([1]) *In L^*, let $\Gamma \subset F(S), B \in F(S), \varepsilon > 0$. If*

$$inf\{H(D(\Gamma), D(\Sigma)) | \Sigma \subset F(S), B \in D(\Sigma)\} < \varepsilon, \tag{10}$$

then B is called a conclusion of Γ of III-type with error less than ε, denoted by $B \in D^3_{L^*,\varepsilon}(\Gamma)$.

Proposition 6. *In L^*, assume that $\Gamma \subset F(S), B \in F(S), \varepsilon > 0$. Let φ be the reflexive transformation of $F(S)$ as defined in Definition 6. Then $B \in D^3_{L^*,\varepsilon}(\Gamma)$ if and only if $\varphi(B) \in D^3_{L^*,\varepsilon}(\varphi(\Gamma))$.*

Proof. $\forall \Sigma \subset F(S), B \in D(\Sigma) \Longleftrightarrow \varphi(B) \in D(\varphi(\Sigma))$. Hence, $inf\{H(D(\Gamma), D(\Sigma)) | \Sigma \subset F(S), B \in D(\Sigma)\} < \varepsilon \Longleftrightarrow inf\{H(D(\varphi(\Gamma)), D(\varphi(\Sigma))) | \varphi(\Sigma) \subset F(S), \varphi(B) \in D(\varphi(\Sigma))\} < \varepsilon$. Let $\varphi(\Sigma) = \Sigma_1$, then

$B \in D^3_{L^*,\varepsilon}(\Gamma) \Longleftrightarrow inf\{H(D(\Gamma), D(\Sigma)) | \Sigma \subset F(S), B \in D(\Sigma)\} < \varepsilon \Longleftrightarrow inf\{H(D(\varphi(\Gamma)), D(\Sigma_1)) | \Sigma_1 \subset F(S), \varphi(B) \in D(\Sigma_1)\} < \varepsilon \Longleftrightarrow \varphi(B) \in D^3_{L^*,\varepsilon}(\varphi(\Gamma))$.

6 Reflexive Transformation φ^* on L^*−Lindenbaum Algebra and Its Properties

Definition 14. ([16])*Let $F(S)$ be the free algebra of the type $(\neg, \vee, \rightarrow)$ generated by the set of atomic formulae $S = \{p_1, \cdots, p_n, \cdots\}$, and $\approx$ be the logical equivalence relation on $F(S)$, then $\approx$ is a congruence relation of type $(\neg, \vee, \rightarrow)$. And $F(S)/\approx$ is called L^*−Lindenbaum algebra, denoted by $[F]$. $\forall A \in F(S)$, the congruence class of A is denoted by $[A]$.*

Definition 15. ([17])*Let $(F(S), \rho)$ be the logic metric space. Define*

$$\rho^*([A], [B]) = \rho(A, B), \quad A, B \in F(S). \tag{11}$$

Then $([F], \rho^*)$ is a metric space.

Definition 16. ([15])*Let φ be the reflexive transformation of $F(S)$ as defined in Definition 6. Define*

$$\varphi^*([A]) = [\varphi(A)], \quad A \in F(S). \tag{12}$$

Then φ^* is called a reflexive transformation of $[F]$.

Proposition 7. *$\varphi^* : [F] \rightarrow [F]$ is an automorphism of L^*−Lindenbaum algebra.*

Proposition 8. *$\varphi^* : [F] \rightarrow [F]$ is an isometric transformation of $([F], \rho^*)$.*

Theorem 6. *Let $\bar{1}, \bar{0}$ be the tautology and contradiction of $F(S)$ respectively. Then $[\bar{1}]$ and $[\bar{0}]$ are fixed points of φ^*.*

Proof. Since $\bar{1}$ is a tautology of $F(S)$, then $\varphi(\bar{1})$ is a tautology of $F(S)$. Hence $\varphi(\bar{1}) \approx \bar{1}$, i.e. $[\varphi(\bar{1})] = [\bar{1}]$. And hence $\varphi^*([\bar{1}]) = [\varphi(\bar{1})] = [\bar{1}]$, which shows that $[\bar{1}]$ is a fixed point of φ^*.

Similarly, we can prove that $[\bar{0}]$ is also a fixed point of φ^*.

Theorem 7. *Let φ be the reflexive transformation of $F(S)$ as defined in Definition 6. $\forall A \in F(S)$, $[A] \vee \varphi^*([A])$ and $[A] \wedge \varphi^*([A])$ are fixed points of φ^*.*

Proof. Since $\varphi^* : [F] \rightarrow [F]$ is an automorphism of $[F]$ and $\varphi^*(\varphi^*([A])) = [A]$, then
$\varphi^*([A] \vee \varphi^*([A])) = \varphi^*([A]) \vee \varphi^*(\varphi^*([A])) = \varphi^*([A]) \vee [A] = [A] \vee \varphi^*([A])$ and
$\varphi^*([A] \wedge \varphi^*([A])) = \varphi^*([A]) \wedge \varphi^*(\varphi^*([A])) = \varphi^*([A]) \wedge [A] = [A] \wedge \varphi^*([A])$
Hence $[A] \vee \varphi^*([A])$ and $[A] \wedge \varphi^*([A])$ are fixed points of φ^*.

Remark 2. In classical logic L, the quotient algebra $F(S)/\approx$ is a Boole algebra, which is also called Lindenbaum algebra and denoted by $[F(S)]$. If φ^* is a reflexive transformation of $[F(S)]$ and $[B]$ is a fixed point of φ^*, then there exists a formula $A \in F(S)$ ($[A]$is not a fixed point of φ^*) such that $[B] = [A] \vee \varphi^*([A])$ or $[B] = [A] \wedge \varphi^*([A])$ (Theorem 6 in paper[15]).

However, this conclusion is not valid in $[F]$. For example, let $B = p_1 \to p_1$, then B is a tautology of $F(S)$ and therefore $[B]$ is a fixed point of φ^*. But $p_1 \vee \neg p_1$ and $p_1 \wedge \neg p_1$ are not tautologies, therefore $[B] \neq [p_1] \vee \varphi^*([p_1])$ and $[B] \neq [p_1] \wedge \varphi^*([p_1])$.

Theorem 8. *Let φ_1^* be a reflexive transformation of $[F(S)]$, and φ_2^* a reflexive transformation of $[F]$. $\forall A \in F(S)$, if $[A]$ is a fixed point of φ_2^*, then $[A]$ is also a fixed point of φ_1^*. Conversely, if $[A]$ is a fixed point of φ_1^*, then there exists a formula $B \in F(S)$ such that $A \approx B$ holds in classical logic, and in addition, $[B]$ is a fixed point of φ_2^*.*

Proof. $\forall A \in F(S)$, if $[A]$ is a fixed point of φ_2^*, then $\varphi_2^*([A]) = [A]$. Hence $[\varphi(A)] = [A]$. i.e. $\varphi(A) \approx A$ in the logic system L^*. Therefore $\varphi(A) \approx A$ holds in classical logic system L, which immediately entails that $\varphi_1^*([A]) = [\varphi(A)] = [A]$. i.e., $[A]$ is a fixed point of φ_1^*.

Conversely, if $[A]$ is a fixed point of φ_1^*, then there exists a formula $A_1 \in F(S)$ such that $[A] = [A_1] \vee \varphi_1^*([A_1]) = [A_1 \vee \varphi(A_1)]$. Let $B = A_1 \vee \varphi(A_1)$, it can be easily verified that $A \approx B$ and $\varphi_2^*([B]) = \varphi_2^*([A_1 \vee \varphi(A_1)]) = [\varphi(A_1 \vee \varphi(A_1))] = [A_1 \vee \varphi(A_1)] = [B]$. Hence $[B]$ is a fixed point of φ_2^*.

Remark 3. In classical logic L, let φ_1^* be a reflexive transformation of $[F(S)]$, and Γ_1 be a set of all fixed points of φ_1^*. In L^*, let φ_2^* be a reflexive transformation of $[F]$, and Γ_2 be a set of all fixed points of φ_2^*. From Theorem 8, we have $\|[\Gamma_1]\| \leq \|[\Gamma_2]\|$ ($\|[\Gamma]\|$ is the cardinal number of set $[\Gamma]$).

Proposition 9. *Let φ_1^* be a reflexive transformation of $[F(S)]$, and φ_2^* be a reflexive transformation of $[F]$. Then the fixed points of φ_1^* (or φ_2^*) exist in pair.*

Proof. Let $[A_1]$ be a fixed point of φ_1^*. Then $\varphi(A_1) \approx A_1$, hence $\neg\varphi(A_1) \approx \neg A_1$. Since φ is a homomorphism of $F(S)$, then $\varphi(\neg A_1) \approx \neg A_1$. Hence $[\neg A_1]$ is fixed point of φ_1^*. i.e. The fixed points of φ_1^* exist in pair.

Similarly, we can show that the fixed points of φ_2^* exist in pair.

Theorem 9. *In classical logic L, let φ_1^* be a reflexive transformation of $[F(S)]$, $[\Gamma_1]$ be a set of all fixed points of φ_1^*, then $[\Gamma_1]$ contains countably many fixed points.*

Proof. Assume that $[\Gamma_1]$ contains a finite number of fixed points. From Proposition 9, we know that the fixed points of φ_1^* exist in pair. Except $[\bar{1}], [\bar{0}]$, $[\Gamma_1]$ contains an even number of fixed points, denoted by $[A_1], \cdots, [A_n], [\neg A_1], \cdots, [\neg A_n]$, and there exist at least two different evaluations $\bar{v_1}, \bar{v_2}$ such that $\bar{v_1}(A_i) = \bar{v_2}(A_i) = 1 (i = 1, \cdots, n)$(This is true, if not, we can exchange $[A_i]$ and $[\neg A_i]$). Then there exists an evaluation v such that $v(A_1) = \cdots = v(A_n) = 1, v(\neg A_1) = \cdots = v(\neg A_n) = 0$, and an evaluation v_i different from v such that $v_i(A_i) = 1 (i = 1, \cdots, n)$.

Let $B = A_1 \wedge \cdots \wedge A_n$. Then $v(B) = 1$. Since $A_i (i = 1, \cdots, n)$ is not a tautology, then B is neither a contradiction nor a tautology. Since $v(\neg A_1) = \cdots =$

$v(\neg A_n) = 0$, then each $\neg A_i (i = 1, \cdots, n)$is not logically equivalent with B. Since there exists an evaluation v_i different from v such that $v_i(A_i) = 1(i = 1, \cdots, n)$, there exists a natural number $j(j \neq i)$ such that $v_i(A_j) = 0$(if not, then $v_i(A_1) = \cdots = v_i(A_n) = 1, v_i(\neg A_1) = \cdots = v_i(\neg A_n) = 0$, hence $v_i = v$), which implies that $v_i(B) = 0$ and therefore each $A_i (i = 1, \cdots, n)$ is not logically equivalent to B. Hence $[B] \neq [\bar{1}], [\bar{0}], [A_1], \cdots, [A_n], [\neg A_1], \cdots, [\neg A_n]$, and $\varphi_1^*([B]) = [\varphi(A_1 \wedge \cdots \wedge A_n)] = [\varphi(A_1) \wedge \cdots \wedge \varphi(A_n)] = [A_1 \wedge \cdots \wedge A_n] = [B]$. i.e., $[B]$is a fixed point of φ_1^*, which contradicts the assumption. Therefore $[\Gamma_1]$ contains countably many fixed points.

Corollary 3. *In logic L^*, let φ_2^* be a reflexive transformation of $[F]$, $[\Gamma_2]$be a set of all fixed points of φ_2^*, then $[\Gamma_2]$ contains countably many fixed points.*

7 Conclusion

In this paper, a reflexive transformation φ in the logic system L^* is initially proposed, and then three different approximate reasoning patterns in system L^* are studied. It is proved that for any $A \in F(S)$ and $\Gamma \subset F(S)$, A is an approximate conclusion of Γ if and only if $\varphi(A)$ is that of $\varphi(\Gamma)$ with the same error. Finally, the cardinal numbers of the sets of fixed points $[\Gamma_1], [\Gamma_2]$ are computed.

Acknowledgements. This work described here is partially supported by the grants from the National Natural Science Foundation of China (NSFC No. 10771129).

References

1. Wang, G.J.: Introduction to mathematical logic and resolution principle, 2nd edn. Science Press, Beijing (2006) (in Chinese)
2. Gottwald, S.: A treatise on many-valued logics, pp. 419–421. Research Studies Press, Baldock (2001)
3. Pei, D.W.: A survey of ten years' studies on fuzzy logic and fuzzy reasoning. Chinese Journal of Engineering Mathematics 21(2), 249–258 (2004) (in Chinese)
4. Wang, Y.Y.: The logic in computer science. Science Press, Beijing (1989) (in Chinese)
5. Cao, Z.L., Wei, C.F.: The application and improvement of modern computer. Science Press, Beijing (2002) (in Chinese)
6. Wang, G.J., Fu, L., Song, J.S.: On the theory of truth degree of propositions in two-valued propositional logic. Science in China series A 31(11), 998–1008 (2001) (in Chinese)
7. Wang, G.J., Leung, Y.: Integrated semantics and logic metric spaces. Fuzzy Sets and Systems 136(1), 71–91 (2003)
8. Wang, G.J., Li, B.J.: Truth degree theory of formulas in Lukasiewicz n-valued propositional logic and limitation theorem. Science in China series E 35(6), 561–569 (2005) (in Chinese)
9. Wang, G.J., Wang, W.: Logical metric spaces. Acta Mathematic Sinica 44(1), 159–168 (2001) (in Chinese)

10. Wang, G.J., Zhang, W.X.: Consistency degrees of finite theories in Lukasiewicz propositional fuzzy logic. Fuzzy Sets and Systems 149, 275–284 (2005)
11. Wang, G.J.: Quantitative logic(I). Chinese Journal of Engineering Mathematics 23(2), 191–215 (2006) (in Chinese)
12. Wang, G.J., She, Y.H.: Topological description of divergency and consistency of two-valued propositional theories. Acta Mathematic Sinica 50(4), 1–10 (2007) (in Chinese)
13. Wang, G.J., Wang, W., Song, S.J.: Topology on the set of maximal consistent propositional theories and the Cantor ternary set. Journal of Shaanxi Normal University (Natural Science Edition) 35(2), 1–5 (2007) (in Chinese)
14. Zhou, H.J., Wang, G.J.: Characterizations of maximal consistent theories in the formal deductive system L^*(NM-logic) and Cantor Space. Fuzzy Sets and Systems 158(23), 2591–2604 (2007)
15. Hu, M.D., Wang, G.J.: Reflexive transformation in classical logic metric space. Journal of Shaanxi Normal University(Natural Science Edition) 37(6), 1–6 (2009) (in Chinese)
16. Wang, G.J.: Non-classical Mathematical Logic and Approximate Reasoning, 2nd edn. Science Press, Beijing (2008) (in Chinese)
17. Wang, W., Wang, G.J.: Continuity of basic operations on pseudo-metric (L)*-Lindenbaum algebra. Journal of Shaanxi Normal University (Natural Science Edition) 33(2), 1–4 (2005) (in Chinese)

A Quantitative Analysis of Rough Logic

Yan-hong She[1] and Guo-jun Wang[2]

[1] College of Sciences, Xi'an Shiyou University, Xi'an 710065, China
e-mail: yanhongshe@gmail.com
[2] College of Mathematics and Information Science, Shaanxi Normal University, Xi'an 710062, China
e-mail: gjwang@snnu.edu.cn

Abstract. The present paper mainly concerns the quantitative analysis of rough logic PRL [1], which is a propositional logic system for rough sets with pre-rough algebra semantics. The concepts of rough(upper, lower) truth degrees on the set of logic formulas in PRL are initially introduced. Then, by grading the rough equality relations, the concepts of rough(upper, lower) similarity degrees are proposed. Finally, three different pseudo-metrics on the set of logic formulas in PRL are obtained, and the corresponding approximate reasoning mechanisms reflecting the idea of rough approximations are established.

Keywords: Rough(upper, lower) truth degree, Rough(upper, lower) similarity degree, Rough(upper, lower) pseudo-metric, Approximate reasoning.

1 Introduction

Rough set theory [2, 3] is proposed by Pawlak to account for the definability of a concept in terms of some elementary ones in an approximation space. It captures and formalizes the basic phenomenon of information granulation. The finer the granulation is, the more concepts are definable in it. For those concepts not definable in an approximation space, the lower and upper approximations for them can be defined. In recent years, as an effective tool in extracting knowledge from data tables, rough set theory has been widely applied in intelligent data analysis, decision making, machine learning and other related fields [4, 5, 6].

As is well known, set theory and logic systems are strongly coupled in the development of modern logic. Since the inception of rough set theory, many scholars have been trying to establish some rough logics corresponding to rough set semantics. The notion of rough logic was initially proposed by Pawlak in [7], in which five rough values, i.e., true, false, roughly true, roughly false and roughly inconsistent were also introduced. This work was

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 145–156.
springerlink.com

subsequently followed by E. Ortowska and Vakarelov in a sequence of papers [8, 9, 10]. There are of course other research works along this research line. For instance, those reported in [1,11,12,13,14,15,16]. Among these research works, PRL [1], a formal logic system corresponding to pre-rough algebra, serves as an interesting formalization of rough sets in the sense that it includes axioms and inference rules, and moreover, it is sound and complete with respect to rough set semantics.

The language of PRL consists of propositional variables(also called atomic formulas) $p_1, p_2, \cdots, p_n, \cdots$, and three primitive logical connectives $\neg$, $\sqcap$ and L. The set of all the logic formulas in PRL is denoted by $F(S)$, which can be formed in the usual manner.

In PRL, three additional logic connectives $\sqcup$, M and $\rightarrow$ are defined as follows: $\forall A, B \in F(S)$,

$$A \sqcup B = \neg (\neg A \sqcap \neg B), \tag{1}$$

$$MA = \neg L \neg A, \tag{2}$$

$$A \rightarrow B = (\neg LA \sqcup LB) \wedge (\neg MA \sqcup MB). \tag{3}$$

A model [16] for PRL (or PRL-model for short) is of the form $M = (U, R, v)$, where the departure from the S5-semantics lies in the meaning function v with respect to the connectives of conjunction $\sqcap$. For $A, B \in F(S)$,

$$v(A \sqcap B) = v(A) \sqcap v(B), \tag{4}$$

where $v(A) \sqcap v(B) = (v(A) \cap v(B)) \cup (v(A) \cap \bar{R}(v(B)) \cap (\bar{R}(v(A) \cap v(B))^c))$ with $\bar{R}$, $\cap$ and c denoting the rough upper approximator [2, 3], usual set-theoretical intersection and set-theoretical complement, respectively.

Definition of truth of a formula A in a model remain the same. Note that for $X, Y \subseteq U, X \sqcap Y$ does not coincide with $X \cap Y$ generally, except one of X, Y is an exact set in the approximation space (U, R). It is also observed that PRL is complete with respect to all these models.

Given two logic formulas p and q, we see that $p \rightarrow p$ is always true and $\neg (p \rightarrow p)$ is always false. Hence we can conclude that $p \rightarrow p$ is the good formula and the $\neg (p \rightarrow p)$ is the bad one. Then one natural question arises: what are the goodness of $p \sqcap q$ and $\neg p \rightarrow Lq$? Until now, several ways(see [17-19]) have been proposed to solve this problem in several commonly used propositional logic systems. However, seen from the viewpoint of rough set theory, these mentioned methods have their own shortcomings because they don't embody the idea of rough approximations. Take the rough formula $p \sqcup \neg p$ as an example. It is not difficult to see that $p \sqcup \neg p$ is not always true in each PRL-model $M = (U, R, v)$, and therefore, its truth degree under previous method in [17,18,19] is strictly less than 1. However, from the viewpoint of rough set theory, it still can be treated as the good formula, because $p \sqcup \neg p$ is roughly true in each PRL-model $M = (U, R, v)$ in the sense of $\bar{R}(v(p \sqcup \neg p)) = U$, in other words, every object $x \in U$ is "possibly" contained in $v(p \sqcup \neg p)$ in each Kripke model $M = (U, R, v)$. Due to this, we need other more plausible measures to evaluate the rough truth degree of formulas in PRL.

In this paper, to evaluate the rough goodness of logic formulas in rough logic PRL, the concepts rough(upper, lower) truth degree are introduced on the set of logic formulas in PRL. Then, based on these fundamental concepts, the concepts of rough(upper, lower) similarity degree are also proposed and their basic properties are investigated. Finally, three different pseudo-metrics are introduced on the set of logic formulas and the corresponding approximate reasoning mechanisms are established.

2 Pre-rough Algebra and Pre-rough Logic PRL

Let's briefly review the basic notions of rough set theory initially proposed by Pawlak [2, 3].

Definition 1. *An approximation is a tuple $AS = (U, R)$, where U is a non-empty set, also called the universe of discourse, R is an equivalence relation on U, representing indiscernibility at the object level.*

Definition 2. *Let $AS = (U, R)$ be an approximation space defined as above. For any set $X \subseteq U$, if X is a union of some equivalence classes produced by R, then we call X a definable set, and otherwise, a rough set. As for rough set X, two definable sets are employed to approximate it from above and from below, respectively. They are*

$$\underline{X} = \{x \in U | [x] \cap \subseteq X\}, \tag{5}$$

$$\overline{X} = \{x | [x] \cap X \neq \emptyset\}, \tag{6}$$

where $[x]$ denotes the equivalence block containing x.

Then we call $\overline{X}(\underline{X})$ rough upper(lower) approximation of X. We will note that X is a definable set if and only if $\overline{X} = \underline{X}$, and therefore, we also treat definable sets as special cases of rough sets.

Definition 3. *A structure $\mathcal{P} = (P, \leq, \sqcap, \sqcup, \neg, L, \rightarrow, 0, 1)$ is a pre-rough algebra, if and only if*

1) $(P, \leq, \sqcap, \sqcup, \rightarrow, 0, 1)$ is a bounded distributive lattice with least element 0 and largest element 1,

2) $\neg\neg a = a$,

3) $\neg (a \sqcup b) = \neg a \sqcap \neg b$,

4) $La \leq a$,

5) $L(a \sqcap b) = La \sqcap Lb$,

6) $LLa = La$,

7) $L1 = 1$,

8) $MLa = La$,

9) $\neg La \sqcup La = 1$,
10) $L(a \sqcup b) = La \sqcup Lb$,
11) $La \leq Lb, Ma \leq Mb$ *imply* $a \leq b$,
12) $a \to b = (\neg La \sqcup Lb) \sqcap (\neg Ma \sqcup Mb)$.

Example 1. Let $\mathbf{3} = (\{0, \frac{1}{2}, 1\}, \leq, \sqcap, \sqcup, \neg, L, \to, 0, 1)$, where $\leq$ is the usual order on real numbers, i.e., $0 \leq \frac{1}{2} \leq 1$, $\sqcap$ and $\sqcup$ are maximum and minimum, respectively. In addition, $\neg 0 = 1, \neg \frac{1}{2} = \frac{1}{2}, \neg 1 = 0$, $L0 = L\frac{1}{2} = 0, L1 = 1$. Then it can be easily checked that **3** is a pre-rough algebra, and also a smallest non-trival pre-rough algebra.

Example 2. Assume that $AS = (U, R)$ is an approximation space, and $P = \{(\underline{X}, \overline{X}) | X \subseteq U\}$. Define operations $\sqcup, \sqcap, \neg, L$ on P as follows: $\forall (\underline{X}, \overline{X}), (\underline{Y}, \overline{Y}) \in P$,

$$(\underline{X}, \overline{X}) \sqcup (\underline{Y}, \overline{Y})) = (\underline{X} \cup \underline{Y}, \overline{X} \cup \overline{Y}), \tag{7}$$

$$(\underline{X}, \overline{X}) \sqcap (\underline{Y}, \overline{Y})) = (\underline{X} \cap \underline{Y}, \overline{X} \cap \overline{Y}), \tag{8}$$

$$\neg (\underline{X}, \overline{X}) = (\neg \overline{X}, \neg \underline{X}), \tag{9}$$

$$L(\underline{X}, \overline{X}) = (\underline{X}, \underline{X}). \tag{10}$$

It can be easily checked that P is closed under the above operations, and moreover, $(P, \sqcup, \sqcap, \neg, L, (\emptyset, \emptyset), (U, U))$ forms a pre-rough algebra.

Definition 4. *The axioms of PRL consist of the formulas of the following form:*
1) $A \to A$,
2) $\neg\neg A \to A$,
3) $A \to \neg\neg A$,
4) $A \sqcap B \to A$,
5)$A \sqcap B \to B \sqcap A$,
6) $A \sqcap (B \sqcup C) \to (A \sqcap B) \sqcup (A \sqcap C)$,
7) $(A \sqcap B) \sqcup (A \sqcap C) \to A \sqcap (B \sqcup C)$,
8) $LA \to A$,
9) $L(A \sqcap B) \to LA \sqcap LB$,
10) $LA \sqcap LB \to L(A \sqcap B)$,
11) $LA \to LLA$,
12) $MLA \to LA$,
13) $L(A \sqcup B) \to LA \sqcup LB$.
The inference rules are as follows:
1) MP rule: $\{A, A \to B\} \vdash B$,
2) HS rule: $\{A \to B, B \to C\} \vdash A \to C$,
3) $\{A\} \vdash B \to A$,
4) $\{A \to B\} \vdash \neg B \to \neg A$,
5) $\{A \to B, A \to C\} \vdash A \to B \sqcap C$,
6) $\{A \to B, B \to A, C \to D, D \to C\} \vdash (A \to C) \to (B \to D)$,

7) $\{A \to B\} \vdash LA \to LB$,
8) $\{A\} \vdash LA$,
9) $\{LA \to LB, MA \to MB\} \vdash A \to B$.

The syntactic notions in PRL, such as theorem and Γ-consequece can be defined in a similar way as in commonly used propositional logic. Here we still use the same denoting symbols such as $\vdash A$, $\Gamma \vdash A$, etc.

As in the case of S5, a model for PRL (or briefly PRL-model) is of the form $M = (U, R, v)$, where the departure from the S5-semantics lies in the definition of the meaning function v with respect to connectives of conjunction $\sqcap$ and implication $\to$. For any $A, B \in F(S)$,

$$v(A \sqcap B) = v(A) \sqcap v(B), \tag{11}$$

$$v(A \to B) = v((\neg LA \sqcup LB) \sqcap (\neg MA \sqcup MB)). \tag{12}$$

Definition of truth of any formula in $M = (U, R, v)$ remains the same: this is if and only if $v(A) = U$. It may then be noticed that $\to$ reflects the rough inclusion: a formula $A \to B$ is true in $M = (U, R, v)$ provided $v(A)$ is roughly included in $v(B)$. Furthermore, $\sqcap/\sqcup$ are operations that reduce to ordinary set intersection/union only when working on definable sets.

A is valid(denoted by $\models A$) if and only if A is true in each PRL-model.

PRL is observed to be complete with respect to the above Kriple semantics. That is,

Theorem 1. $\vdash A$ *if and only if* $\models A$.

3 Rough(Upper, Lower) Truth Degrees of Rough Formulas

Presented below is the quantitative theory of rough logic PRL in any given model. As shown below, our discussion is based on finite PRL-models.

Definition 5. *Let* $A \in F(S)$ *and* $\mathcal{M}= (U, R, v)$ *be a finite PRL-model. Define*

$$\tau_{\mathcal{M}}(A) = \frac{|v(A)|}{|U|}, \tag{13}$$

$$\bar{\tau}_{\mathcal{M}}(A) = \frac{|\overline{v(A)}|}{|U|}, \tag{14}$$

$$\underline{\tau}_{\mathcal{M}}(A) = \frac{|\underline{v(A)}|}{|U|}. \tag{15}$$

Then we call $\tau_{\mathcal{M}}(A)$, $\bar{\tau}_{\mathcal{M}}(A)$ *and* $\underline{\tau}_{\mathcal{M}}(A)$ *the rough truth degree, rough upper truth degree and rough lower truth degree, respectively, in the given PRL-model* $\mathcal{M}= (U, R, v)$.

Example 3. Let $\mathcal{M}= (U, R, v)$ be a PRL-model, where $U = \{1, 2, 3, 4, 5\}$, $R = \{(1, 1), (1, 2), (2, 1), (2, 2), (3, 3), (3, 4), (4, 3), (4, 4), (5, 5)\}$ and $v(p) = \{1, 3\}, v(q) = \{2, 5\}$. Compute the rough truth degree, rough upper truth degree and rough lower truth degree of $p \sqcap q$.

Solution: $\tau_{\mathcal{M}}(p \sqcap q) = \frac{|v(p \sqcap q)|}{|U|} = \frac{|v(p) \sqcap v(q)|}{|U|} = \frac{|(v(p) \cap v(q)) \cup (v(p) \cap \overline{v(q)} \cap \overline{v(p) \cap v(q)}^{c})|}{|U|} = \frac{|\{1,3\} \cap \{1,2,5\}|}{5} = \frac{1}{5}$.

$\bar{\tau}_{\mathcal{M}}(p \sqcap q) = \frac{|\overline{v(p \sqcap q)}|}{|U|} = \frac{|\overline{v(p)} \cap \overline{v(q)}|}{|U|} = \frac{|\{1,2,3,4\} \cap \{1,2,5\}|}{5} = \frac{2}{5}$.

$\underline{\tau}_{\mathcal{M}}(p \sqcap q) = \frac{|\underline{v(p \sqcap q)}|}{|U|} = \frac{|(\underline{v(p)} \cap \underline{v(q)}|}{|U|} = \frac{|\emptyset|}{5} = 0$.

The concepts of rough(upper,lower) truth degrees enjoy the following properties.

Proposition 1. *Let $\mathcal{M}= (U, R, v)$ be a finite PRL-model, and $\tau_{\mathcal{M}}(\bar{\tau}_{\mathcal{M}}, \underline{\tau}_{\mathcal{M}})$ be the rough(upper, lower) truth degree mappings defined as in (13-15). Then $\forall A, B \in F(S)$,*

1) $0 \leq \underline{\tau}_{\mathcal{M}}(A) \leq \tau_{\mathcal{M}}(A) \leq \bar{\tau}_{\mathcal{M}}(A) \leq 1$.

2) $\tau_{\mathcal{M}}(A) = 1$ if and only if $v(A) = U$, i.e., A is true in the PRL-model $\mathcal{M}= (U, R, v)$.

$\bar{\tau}_{\mathcal{M}}(A) = 1$ if and only if $\overline{v(A)} = U$, i.e., $v(A)$ is roughly true in $\mathcal{M}= (U, R, v)$.

3) $\tau_{\mathcal{M}}(\neg A) = 1 - \tau_{\mathcal{M}}(A)$, $\bar{\tau}_{\mathcal{M}}(\neg A) = 1 - \underline{\tau}_{\mathcal{M}}(A)$.

4) $\bar{\tau}^{R}_{\mathcal{M}}(A) = \bar{\tau}^{R}_{\mathcal{M}}(B)$ if and only if $\overline{v(A)} = \overline{v(B)}$,

$\underline{\tau}^{R}_{\mathcal{M}}(A) = \underline{\tau}^{R}_{\mathcal{M}}(B)$ if and only if $\underline{v(A)} = \underline{v(B)}$.

5) $\vdash A \to B$ implies that $\bar{\tau}_{\mathcal{M}}(A) \leq \bar{\tau}_{\mathcal{M}}(B)$ and $\underline{\tau}_{\mathcal{M}}(A) \leq \underline{\tau}_{\mathcal{M}}(B)$;

$\vdash A \leftrightarrow B$ implies that $\bar{\tau}_{\mathcal{M}}(A) = \bar{\tau}_{\mathcal{M}}(B), \underline{\tau}_{\mathcal{M}}(A) = \underline{\tau}_{\mathcal{M}}(B)$.

Definition 6. *Let $A \in F(S)$ and $\mathcal{M}= (U, R, v)$ be a finite PRL-model. Define*

$$\xi_{\mathcal{M}}(A, B) = \tau_{\mathcal{M}}((A \to B) \sqcap (A \to B)), \tag{16}$$

$$\bar{\xi}_{\mathcal{M}}(A, B) = \tau_{\mathcal{M}}((MA \to MB) \sqcap (MB \to MA)), \tag{17}$$

$$\underline{\xi}_{\mathcal{M}}(A, B) = \tau_{\mathcal{M}}((LA \to LB) \sqcap (LB \to LA)). \tag{18}$$

Then we call $\xi_{\mathcal{M}}(A, B)$, $\bar{\xi}_{\mathcal{M}}(A, B)$ and $\underline{\xi}_{\mathcal{M}}(A,B)$ the rough similarity degree, the rough upper similarity degree and the rough lower similarity degree between A and B, respectively, in the PRL-model $\mathcal{M}= (U, R, v)$, and we also call $\xi_{\mathcal{M}}$, $\bar{\xi}_{\mathcal{M}}$ and $\underline{\xi}_{\mathcal{M}}$ the rough similarity degree mapping, rough upper similarity degree mapping and rough lower similarity degree mapping, respectively.

Rough(upper, lower) similarity degree mappings enjoy the following properties.

Proposition 2. *Let $\mathcal{M}= (U, R, v)$ be a finite PRL-model, and $\xi_{\mathcal{M}}(\bar{\xi}_{\mathcal{M}}, \underline{\xi}_{\mathcal{M}})$ be the rough(upper, lower) truth degree mappings defined as above. Then $\forall A, B \in F(S)$,*

1) $0 \leq \xi_{\mathcal{M}}(A,B), \bar{\xi}_{\mathcal{M}}(A,B), \underline{\xi}_{\mathcal{M}}(A,B) \leq 1$.
2)$\xi_{\mathcal{M}}(A,B) = \xi_{\mathcal{M}}(B,A), \bar{\xi}^R_{\mathcal{M}}(A,B) = \bar{\xi}^R_{\mathcal{M}}(B,A), \underline{\xi}^R_{\mathcal{M}}(A,B) = \underline{\xi}^R_{\mathcal{M}}(B,A)$.
3) $\bar{\xi}^R_{\mathcal{M}}(A,B) = \xi^R_{\mathcal{M}}(MA, MB), \underline{\xi}^R_{\mathcal{M}}(A,B) = \xi^R_{\mathcal{M}}(LA, LB)$.
4) $\bar{\xi}^R_{\mathcal{M}}(A,B) = 1$ *if and only if* $\overline{v(A)} = \overline{v(B)}$,
$\underline{\xi}^R_{\mathcal{M}}(A,B) = 1$ *if and only if* $\underline{v(A)} = \underline{v(B)}$,
$\xi_{\mathcal{M}}(A,B) = 1$ *if and only if* $\overline{v(A)} = \overline{v(B)}, \underline{v(A)} = \underline{v(B)}$.
5)$\vdash A \leftrightarrow B$ *entails that* $\bar{\xi}^R_{\mathcal{M}}(A,B) = \underline{\xi}^R_{\mathcal{M}}(A,B) = \xi_{\mathcal{M}}(A,B) = 1$.

Proof. 1), 2) and 3) are evident from the definition of rough(upper, lower) truth degrees.

$$
\begin{aligned}
4) \quad \bar{\xi}^R_{\mathcal{M}}(A,B) = 1 &\Leftrightarrow \tau_{\mathcal{M}}((MA \to MB) \sqcap (MB \to MA)) = 1 \\
&\Leftrightarrow \tfrac{|v((MA \to MB) \sqcap (MB \to MA))|}{|U|} = 1 \\
&\Leftrightarrow \tfrac{|(v(MA)^c \cup v(MB)) \cap (v(MB)^c \cup v(MA))|}{|U|} = 1 \\
&\Leftrightarrow (v(MA)^c \cup v(MB)) \cap (v(MB)^c \cup v(MA)) = U \\
&\Leftrightarrow (\overline{v(A)}^c \cup \overline{v(B)}) \cap (\overline{v(B)}^c \cup \overline{v(A)}) = U \\
&\Leftrightarrow \overline{v(A)} \subseteq \overline{v(B)}, \overline{v(B)} \subseteq \overline{v(A)} \\
&\Leftrightarrow \overline{v(A)} = \overline{v(B)}.
\end{aligned}
$$

The other two conclusions can be proved in a similar way, and hence are omitted here.

5) If $\vdash A \leftrightarrow B$, then we have from Theorem 1 that in the given PRL-model $\mathcal{M}= (U, R, v)$, $v(A \leftrightarrow B) = v((A \to B) \sqcap (B \to A)) = U$, i.e., $(\underline{v(A)}^c \cup \underline{v(B)}) \cap (\overline{v(A)}^c \cup \overline{v(B)}) \cap (\underline{v(B)}^c \cup \underline{v(A)}) \cap (\overline{v(B)}^c \cup \overline{v(A)}) = U$, which immediately entails that $\overline{v(A)} = \overline{v(B)}, \underline{v(A)} = \underline{v(B)}$.

By introducing a pseudo-metric on the set of logic formulae in PRL, one can establish the corresponding approximate reasoning mechanism by means of the pseudo-metric. As shown below, such pseudo-metrics are induced by rough (upper, lower) similarities.

Definition 7. *Let* $\mathcal{M}= (U, R, v)$ *be a finite Kripke model. Define three nonnegative mappings* $\rho_{\mathcal{M}}, \bar{\rho}_{\mathcal{M}}, \underline{\rho}_{\mathcal{M}} : F(S) \times F(S) \longrightarrow [0,1]$ *as follows:*

$$\rho_{\mathcal{M}}(A,B) = 1 - \xi^R_{\mathcal{M}}(A,B), \tag{19}$$

$$\bar{\rho}_{\mathcal{M}}(A,B) = 1 - \bar{\xi}^R_{\mathcal{M}}(A,B), \tag{20}$$

$$\underline{\rho}_{\mathcal{M}}(A,B) = 1 - \underline{\xi}^R_{\mathcal{M}}(A,B). \tag{21}$$

The following proposition states that the above defined mappings, i.e., $\rho_{\mathcal{M}}, \bar{\rho}_{\mathcal{M}}, \underline{\rho}_{\mathcal{M}} : F(S) \times F(S) \longrightarrow [0,1]$, are pseudo-metrics on the set of logic formulae in PRL.

Proposition 3. *Let* $\mathcal{M}= (U, R, v)$ *be a finite Kripke model and* $\rho_{\mathcal{M}}, \bar{\rho}_{\mathcal{M}}, \underline{\rho}_{\mathcal{M}}$ *be the three nonnegative mappings defined above. Then* $\forall A, B, C \in F(S)$,

1) $\rho_{\mathcal{M}}(A,A)=\bar{\rho}_{\mathcal{M}}(A,A)=\underline{\rho}_{\mathcal{M}}(A,A)=0$,
2) $\rho_{\mathcal{M}}(A,B)=\rho_{\mathcal{M}}(B,A), \bar{\rho}_{\mathcal{M}}(A,B)=\bar{\rho}_{\mathcal{M}}(B,A), \underline{\rho}_{\mathcal{M}}(A,B)=\underline{\rho}_{\mathcal{M}}(B,A)$,
3) $\rho_{\mathcal{M}}(A,C)\leq\rho_{\mathcal{M}}(A,B)+\rho_{\mathcal{M}}(B,C)$,
$\bar{\rho}_{\mathcal{M}}(A,C)\leq\bar{\rho}_{\mathcal{M}}(A,B)+\bar{\rho}_{\mathcal{M}}(B,C)$,
$\underline{\rho}_{\mathcal{M}}(A,C)\leq\underline{\rho}_{\mathcal{M}}(A,B)+\underline{\rho}_{\mathcal{M}}(B,C)$.

Proof. Both 1) and 2) are clear from the definition.

3) In what follows, we aim to show that $\rho^R_{\mathcal{M}}(A,C)\leq\rho^R_{\mathcal{M}}(A,B)+\rho^R_{\mathcal{M}}(B,C)$, i.e.,

$$1-\frac{|v((A\to C)\sqcap(C\to A))|}{|U|}$$

$$\leq 1-\frac{|v((A\to B)\sqcap(B\to A))|}{|U|}+1-\frac{|v((B\to C)\sqcap(C\to B))|}{|U|}. \qquad (22)$$

Let $E=v((A\to C)\sqcap(C\to A)), F=v((A\to B)\sqcap(B\to A)), G=v((B\to C)\sqcap(C\to B))$.

Then, to prove (22), i.e., $|U-E|\leq|U-F|+|U-G|$, it suffices to show that $F\cap G\subseteq E$.

It can be easily checked that

$$E=(\underline{v(A)}^c\cup\underline{v(C)})\cap(\overline{v(A)}^c\cup\overline{v(C)})\cap(\underline{v(C)}^c\cup\underline{v(A)})\cap(\overline{v(C)}^c\cup\overline{v(A)}),$$
$$F=(\underline{v(A)}^c\cup\underline{v(B)})\cap(\overline{v(A)}^c\cup\overline{v(B)})\cap(\underline{v(B)}^c\cup\underline{v(A)})\cap(\overline{v(B)}^c\cup\overline{v(A)}),$$
$$G=(\underline{v(B)}^c\cup\underline{v(C)})\cap(\overline{v(B)}^c\cup\overline{v(C)})\cap(\underline{v(C)}^c\cup\underline{v(B)})\cap(\overline{v(C)}^c\cup\overline{v(B)}).$$

Denote by $E_i, F_i, G_i (i=1,2,3,4)$ the ith subset of E, F and G, respectively, from left to right.

$\forall x\in F\cap G$, then $x\in F_i, G_i (i=1,2,3,4)$. There are two cases to be considered below:

Case 1. If $x\in\underline{v(B)}$, then we have from $x\in G_1$ that $x\in\underline{v(C)}$, which yields $x\in E_1$. We can also obtain $x\in E_2$ because of $\underline{v(C)}\subseteq\overline{v(C)}$. Furthermore, we have from $x\in F_3$ that $x\in\underline{v(A)}$, whence $x\in E_3$ and $x\in E_4$ immediately follow. Hence $x\in E_1\cap E_2\cap E_3\cap E_4=E$.

Case 2. If $x\bar{\in}\underline{v(B)}$, then there are still two subcases to consider.

If $x\in\overline{v(B)}$, then we have from $x\in F_1$ that $x\in\underline{v(A)}^c$, which entails $x\in E_1$.

Similarly, we can obtain $x\in\overline{v(C)}$ from $x\in G_2$, which implies that $x\in E_2$. Moreover, it follows from $x\in G_3$ that $x\in\underline{v(C)}^c$, and therefore $x\in E_3$. Also, $x\in F_4$ implies $x\in\overline{v(A)}$, which entails $x\in E_4$. And hence $x\in E_1\cap E_2\cap E_3\cap E_4=E$.

If $x\bar{\in}\overline{v(B)}$, then we obtain $x\in\underline{v(A)}^c$ from $x\in F_1$, which yields $x\in E_1$.

Similarly, we can prove that $x\in E_2$, $x\in E_3$ and $x\in E_4$ by $x\in F_2$, $x\in G_3$ and $x\in G_4$, respectively. Hence $x\in E_1\cap E_2\cap E_3\cap E_4=E$.

This completes the proof of $F \cap G \subseteq E$, whence $U - E \subseteq (U - F) \cup (U - G)$ immediately follows, which entails that $|U - E| \le |(U - F) \cup (U - G)| \le |U - F| + |U - G|$.

On account of Proposition 3, we will call $\rho^R_{\mathcal{M}}, \bar{\rho}^R_{\mathcal{M}}$ and $\underline{\rho}^R_{\mathcal{M}}$ the rough pseudo-metric, rough upper pseudo-metric and rough lower pseudo-lower metric, respectively in the sequel.

In what follows, we aim to define the concepts of (n)-rough truth degree, (n)-rough upper truth degree and (n)-rough lower truth degree, respectively, by collecting all the finite PRL-models indistinguishable to the given formula (explained below) but not in some given PRL-model, and hence it shows more rationality.

Specifically, we will consider the class of PRL-model $\mathcal{M}= (U, R, v)$ satisfying $|U| = n$ below. Particularly, we will only consider the subclass(denoted by $\mathcal{M}_n$) of the PRL-models $\mathcal{M}= (U, R, v)$ satisfying $U = \{1, 2, \cdots, n\}$. Moreover, for any formula $A(p_1, \cdots, p_n) \in F(S)$ and $\mathcal{M}_1= (U, R_1, v_1)$, $\mathcal{M}_2= (U, R_2, v_2) \in \mathcal{M}_n$, we say that $\mathcal{M}_1$ and $\mathcal{M}_2$ are indistinguishable with respect to $A(p_1, \cdots, p_m)$ if and only if $R_1 = R_2$ and $v_1(p_i) = v_2(p_i), i = 1, 2, \cdots, m$. We denote by $\mathcal{M}_{n,A}$(or briefly $\mathcal{M}_n$ if it is clear from the context) the set of PRL-models indistinguishable to A.

Definition 8. $\forall A \in F(S)$, *define*

$$\tau^R_n(A) = \frac{\Sigma\{\tau^R_{\mathcal{M}}(A) | \mathcal{M} \in \mathcal{M}_{n,A}\}}{|\mathcal{M}_{n,A}|}, \tag{23}$$

$$\overline{\tau}^R_n(A) = \frac{\Sigma\{\overline{\tau}^R_{\mathcal{M}}(A) | \mathcal{M} \in \mathcal{M}_{n,A}\}}{|\mathcal{M}_{n,A}|}, \tag{24}$$

$$\underline{\tau}^R_n(A) = \frac{\Sigma\{\underline{\tau}^R_{\mathcal{M}}(A) | \mathcal{M} \in \mathcal{M}_{n,A}\}}{|\mathcal{M}_{n,A}|}, \tag{25}$$

then we call $\tau^R_n(A), \overline{\tau}^R_n(A), \underline{\tau}^R_n(A)$ *the (n)-rough truth degree, (n)-rough upper truth degree and (n)-rough lower truth degree, respectively.*

Example 4. Compute the (3)-rough upper truth degree $\bar{\tau}_3(p)$ and (3)-rough lower truth degree $\underline{\tau}_3(p)$ of p.

Solution: $\forall \mathcal{M}= (U, R, v) \in \mathcal{M}_3$, to compute $\bar{\tau}_3(p), \underline{\tau}_3(p)$, we need to consider all the cases of R and v. It can be easily checked that there are five equivalence relations on the set $\{\{1, 2, 3\}$ in total, they are

$R_1 = \{(1, 1), (2, 2), (3, 3)\}$, $R_2 = \{(1, 1), (2, 2), (3, 3), (1, 2), (2, 1)\}$,
$R_3 = \{(1, 1), (2, 2), (3, 3), (1, 3), (3, 1)\}$, $R_4 = \{(1, 1), (2, 2), (3, 3), (2, 3), (3, 2)\}$,
$R_5 = U \times U$. $v(p)$,

which runs over the powerset of U, has $2^3 = 8$ possible cases. They are $V_1 = \{\emptyset\}$, $V_2 = \{1\}$, $V_3 = \{2\}$, $V_4 = \{3\}$, $V_5 = \{1, 2\}$, $V_6 = \{1, 3\}$, $V_7 = \{2, 3\}$, $V_8 = \{1, 2, 3\}$. We denote by $\mathcal{M}^{i,j}$ the PRL-model corresponding to $R_i, V_j (1 \le i \le 5, 1 \le j \le 8)$, i.e., $\mathcal{M}^{i,j} = (U, R_i, V_j)$ below. Then

$\Sigma\{\bar{\tau}_{\mathcal{M}}(p) \mid \mathcal{M} \in \mathcal{M}^{1,j}, 1 \leq j \leq 8\} = 0 + \frac{1}{3} + \frac{1}{3} + \frac{1}{3} + \frac{2}{3} + \frac{2}{3} + \frac{2}{3} + 1 = 4,$
$\Sigma\{\underline{\tau}_{\mathcal{M}}(p) \mid \mathcal{M} \in \mathcal{M}^{1,j}, 1 \leq j \leq 8\} = 0 + \frac{1}{3} + \frac{1}{3} + \frac{1}{3} + \frac{2}{3} + \frac{2}{3} + \frac{2}{3} + 1 = 4,$
$\Sigma\{\bar{\tau}_{\mathcal{M}}(p) \mid \mathcal{M} \in \mathcal{M}^{2,j}, 1 \leq j \leq 8\} = 0 + \frac{2}{3} + \frac{2}{3} + \frac{1}{3} + \frac{2}{3} + \frac{3}{3} + \frac{3}{3} + \frac{3}{3} = \frac{16}{3},$
$\Sigma\{\underline{\tau}_{\mathcal{M}}(p) \mid \mathcal{M} \in \mathcal{M}^{2,j}, 1 \leq j \leq 8\} = 0 + 0 + 0 + \frac{1}{3} + \frac{2}{3} + 0 + 0 + \frac{3}{3} = 2.$
Similarly, we obtain
$\Sigma\{\bar{\tau}_{\mathcal{M}}(p) \mid \mathcal{M} \in \mathcal{M}^{3,j}, 1 \leq j \leq 8\} = \frac{16}{3}$, $\Sigma\{\underline{\tau}_{\mathcal{M}}(p) \mid \mathcal{M} \in \mathcal{M}^{3,j}, 1 \leq j \leq 8\} = 2,$
$\Sigma\{\bar{\tau}_{\mathcal{M}}(p) \mid \mathcal{M} \in \mathcal{M}^{4,j}, 1 \leq j \leq 8\} = \frac{16}{3}$, $\Sigma\{\underline{\tau}_{\mathcal{M}}(p) \mid \mathcal{M} \in \mathcal{M}^{4,j}, 1 \leq j \leq 8\} = 2,$
$\Sigma\{\bar{\tau}_{\mathcal{M}}(p) \mid \mathcal{M} \in \mathcal{M}^{5,j}, 1 \leq j \leq 8\} = 7$, $\Sigma\{\underline{\tau}_{\mathcal{M}}(p) \mid \mathcal{M} \in \mathcal{M}^{5,j}, 1 \leq j \leq 8\} = 1.$
And hence, $\bar{\tau}_3(p) = \frac{1}{|\mathcal{M}_3|}\Sigma\{\bar{\tau}_{\mathcal{M}}(p) \mid \mathcal{M} \in \mathcal{M}_3\} = \frac{4+3\times\frac{16}{3}+7}{40} = \frac{27}{40}$,
$\underline{\tau}_3(p) = \frac{1}{|\mathcal{M}_3|}\Sigma\{\underline{\tau}_{\mathcal{M}}(p) \mid \mathcal{M} \in \mathcal{M}_3\} = \frac{4+3\times 2+1}{40} = \frac{11}{40}$.

Definition 9. $\forall A, B \in F(S)$, *define*

$$\xi_n(A,B) = \tau_n((A \to B) \sqcap (B \to A)), \tag{26}$$

$$\bar{\xi}_n(A,B) = \tau_n((MA \to MB) \sqcap (MB \to MA)), \tag{27}$$

$$\underline{\xi}_n(A,B) = \tau_n((LA \to LB) \sqcap (LB \to LA)), \tag{28}$$

then we call $\xi_n^R(A,B), \bar{\xi}_n^R(A,B), \underline{\xi}_n^R(A,B)$ *the (n)-rough similarity degree, (n)-rough upper similarity degree and (n)-rough lower similarity degree between* A *and* B*, respectively.*

Definition 10. $\forall A, B \in F(S)$, *define three nonnegative functions* $\rho_n, \overline{\rho}_n, \underline{\rho}_n$: $F(S) \times F(S) \longrightarrow [0,1]$ *as follows:*

$$\rho_n(A,B) = 1 - \xi_n(A,B), \tag{29}$$

$$\overline{\rho}_n(A,B) = 1 - \bar{\xi}_n(A,B), \tag{30}$$

$$\underline{\rho}_n(A,B) = 1 - \underline{\xi}_n(A,B). \tag{31}$$

Proposition 4. $\rho_n, \overline{\rho}_n, \underline{\rho}_n$ *are pseudo-metrics on the set of logic formulae in PRL.*

Proof. It follows immediately from Proposition 3 and Definition 10.

On basis of those pseudo-metrics proposed as above, three different approximate reasoning mechanisms reflecting the idea of rough approximation are presented below.

Definition 11. *Let* $\Gamma \subseteq F(S), \varepsilon > 0$. *If* $\rho_n(A, D(\Gamma)) < \varepsilon$, *then we say that* A *is a rough approximate consequence of* Γ *with error less than* ε.

Similarly, if $\bar{\rho}_n(A, D(\Gamma)) < \varepsilon$, *then we call* A *the rough upper consequence of* Γ *with error less than* ε, *and if* $\underline{\rho}_n(A, D(\Gamma)) < \varepsilon$, *then we call* A *the rough lower consequence of* Γ *with error less than* ε.

4 Concluding Remarks

A quantitative analysis of the rough logic PRL is given in this paper. By grading the concepts of rough truth, the concepts of rough(upper, lower) truth degrees are initially introduced on the set of rough formulas in PRL. Then, based on the fundamental concept of rough truth degree, rough(upper, lower) similarity degrees are also proposed and some of their basic properties are investigated. Finally, three different pseudo-metrics are introduced on the set of formulas in PRL, and the corresponding approximate reasoning mechanisms reflecting the idea of rough approximations are established.

References

1. Berjee, M.: Rough sets through algebraic logic. Fundamenta Informaticae 32, 213–220 (1997)
2. Pawlak, Z.: Rough sets. International Journal of Computer and Information Sciences 11(5), 341–356 (1982)
3. Pawlak, Z.: Rough Sets-Theoretical aspects of reasoning about data. Knowledge Academic Publishers (1991)
4. Lin, T.Y., Cercone, N.: Rough sets and data mining-analysis of imprecise data. Kluwer Academic Publishers, Dordrecht (1997)
5. Polkowski, L., Skowron, A.: Rough sets in knowledge discovery1: methodology and applications. Studies of fuzziness and soft computing, vol. 18. Physica-Verlag, Heidlberg (1998)
6. Polkowski, L., Skowron, A.: Rough sets in knowledge discovery2: applications, case studies and software systems. Studies of fuzziness and soft computing, vol. 19. Physica-Verlag, Heidlberg (1998)
7. Pawlak, Z.: Rough logic. Bull. Polish Acad. Sc. (Tech. Sc.) 35, 253–258 (1987)
8. Orlowska, E.: Kripke semantics for knowledge representation logics. Studia Logica XLIX, 255–272 (1990)
9. Vakarelov, D.: A modal logic for similarity relations in Pawlak knowledge representaion systems. Fundamenta Informaticae 15, 61–79 (1991)
10. Vakarelov, E.: Modal logics for knowledge representation systems. Theoretical Computer Science 90, 433–456 (1991)
11. Duntsch, I.: A logic for rough sets. Theoretical Computer Science 179, 427–436 (1997)
12. Chakrabotry, M.K., Banerjee, M.: Rough consequence. Bull. Polish Acad. Sc. (Math.) 41(4), 299–304 (1993)
13. Dai, J.H.: Logic for rough sets with rough double stone algebraic semantics. In: Ślęzak, D., Wang, G., Szczuka, M.S., Düntsch, I., Yao, Y. (eds.) RSFDGrC 2005. LNCS (LNAI), vol. 3641, pp. 141–148. Springer, Heidelberg (2005)
14. Banerjee, M.: Logic for rough truth. Fundamenta Informaticae 71(2-3), 139–151 (2006)
15. Pagliani, P., Chakraborty, M.: A geometry of approximation. Rough set theory: logic, algebra and topology of conceptual patterns. Springer, Heidelberg (2008)

16. Banerjee, M.: Propositional logic from rough set theory. In: Peters, J.F., Skowron, A., Düntsch, I., Grzymała-Busse, J.W., Orłowska, E., Polkowski, L. (eds.) Transactions on Rough Sets VI. LNCS, vol. 4374, pp. 1–25. Springer, Heidelberg (2007)
17. Wang, G.J., Zhou, H.J.: Introduction to mathematical logic and resolution principle. Science Press, Alpha Science International Limited, Beijing, UK (2009)
18. Wang, G.J., Zhou, H.J.: Quantitative logic. Information Sciences 179, 226–247 (2009)
19. Pavelka, J.: On fuzzy logic I, II, III, Zeitschr. Math. Logik Grund. d. Math. 25, 45–72, 119–134, 447–464 (1979)

Applying Association Rules to ε-Reduction of Finite Theory in Two-Valued Propositional Logic

Li-feng Li[1] and Nan Yang[2]

[1] School of Science, Xi'an University of Posts and Telecommunications, Xi'an 710121, China
e-mail: hlw12-12@sohu.com

[2] School of Computer Science, Xi'an University of Posts and Telecommunications, Xi'an 710121, China
e-mail: yangnanchina@foxmail.com

Abstract. The theory of association rules is an issue in recent years since it has been successfully applied in a wide range of domains, and removing redundant formula in a propositional theory is another issue, but development of the two theory seems independently. In this paper, the mutual relationship between them are investigated by introducing the formal context $(\Omega_\Gamma, \Gamma, I)$, then the theory of ε-reduction of finite theory Γ in two-valued propositional logic is proposed. By introducing the association rules to the formal context $(\Omega_\Gamma, \Gamma, I)$, judgment theorems of ε-consistent theorems are examined, and two approaches to explore ε-reduction are presented.

Keywords: Formal context,association rules, consistent degree, ε-reduction.

1 Introduction

A knowledge base is redundant if it contains parts that can be removed without reducing the information it carries. Removing redundant clause in a formula and redundant formula in a propositional theory are important for some reasons. Firstly, it has some computational advantage in some cases. Moreover, simplifying a formula or theory leads to representation of the same knowledge that is easier to understand, as a large amount of redundancy may obscure the meaning of the represented knowledge. The redundant part of a knowledge base can instead be the core of the knowledge it represents. They are so important that check redundancy of propositional formula and propositional theory are focus these years. The complexity of some problems related to the redundancy of propositional CNF formulae, Horn formulae and Non-monotonic reasoning are studied in [1,2,3], and redundancy of propositional theory in L^* system are studied in [4], all this studies from a pure

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 157–165.
springerlink.com

logic standpoint. Re.[5] re-consider the redundancy of finite propositional theory Γ in two-valued propositional logic by means of concept lattice and obtains many interesting results. Let $D(\Gamma)$ be the set of all Γ conclusions. We called $\Gamma_0 \subseteq \Gamma$ is a reduction of Γ if $D(\Gamma_0) = D(\Gamma)$ and for any $A \in \Gamma_0$, $D(\Gamma) \neq D(\Gamma_0) \setminus \{A\}$, it means that Γ_0 is the minimal irredundant subset of Γ.

Data mining has been extensively addressed for last years as the computational part of Knowledge Discovery in Databases (KDD), specially the problem of discovering association rules. Its aim is to exhibit relationships be- tween item-sets(sets of binary attributes) in large databases. Association rules have been successfully applied in a wide range of domains, among which marketing decision support, diagnosis and medical research support, geographical and statistical data, etc. And various approaches to mining association rules have been proposed[6,7,8].

As we know, in propositional logic, a theory is consistent or inconsistent, there is no intermediate situations, i.e. the concept of consistency of a theory is crisp rather than fuzzy. In order to distinguish two consistent sets Γ and Σ, the concept of consistency degree is introduced in [9,10]. Let $\delta(\Gamma)$ stands for the consistent degree of Γ. In this sense, the way of which elimination redundancy in finite propositional theory in [4] is to explore $\Gamma_0 \subseteq \Gamma$ such that $\delta(\Gamma_0) = \delta(\Gamma)$. We can observe that if we eliminate different formula from Γ_0, $\delta(\Gamma_0 \setminus \{A\})$ may changes. The purpose of this paper is trying to give methods to explore the minimal subset of Γ whose consistent degree is within the consistent degree of Γ plus an error ε. More detail, we firstly give the definition of ε-reduction of theory Γ at the given error ε, then give judgment theorems of ε-reduction of theory Γ and some properties by using the association rules, two methods to explore the ε-reduction and examples are given in Section 4,and Section 5 is conclusion.

2 ε-Reduction of Theory

A formal context (G, M, I) is consists of two sets G and M and a relation I between G and M. The elements of G are called objects and elements of M are called the attributes of the context. In order to express that an object g is in the relation I with an attribute m. we write gIm. And for $A \subseteq G, B \subseteq M$, define the operators $f(A) = \{m \in M \mid gIm \text{ for all } g \in A\}$ and $h(B) = \{g \in G \mid gIm \text{ for all } m \in B\}$.

An implication between attributes (in M) is a pair of subsets of the attribute set M. It is denoted by $P \to Q$, where P is the premise and Q is conclusion.

Definition 1. ($Will.R$[11]) *A subset $T \subseteq M$ respects an implication $P \to Q$ if $P \nsubseteq T$ or $Q \subseteq T$. T respects a set L of implications if T respects every single implication in L. $P \to Q$ holds in a set $\{T_1, T_2, \cdots\}$ of subsets if each of the subsets T_i respects the implication $P \to Q$. $P \to Q$ hods in a context (G, M, I)*

if it holds in the system of object intents. In this case, we also say that $P \to Q$ *is an implication of the context* (G, M, I)*.* P *is the premise of* Q*.*

Theorem 1. $(Will.R$[11]$)$ *An implication* $P \to Q$ *holds in a context* (G, M, I) *if and only if* $Q \subseteq fh(Q)$*.*

What attribute implications represent are deterministic. However, noises and uncertainties are prevailing in realistic domains. It is a must to introduce possibility and probability into knowledge. In addition, to make the discovery knowledge statistically signicant, it is also prerequisite that the knowledge should be support by enough instances (or objects). Therewithal, association rules emerged[6].

Definition 2. *With respect to the support threshold* θ *and confidence threshold* φ*,* $P \to Q$ *is an association rule holding in context* (G, M, I)*, if it satisfies* $\mid h(P \cup Q) \mid \geq \mid G \mid \times \theta$ *and* $\frac{h(P \cup Q)}{h(P)} \geq \varphi$*.*

In two-valued propositional logic, Given a theory $\Gamma = \{A_1, A_2, \cdots, A_n\} \subset F(S)$ contains m-different atomic formulas of S, assume there are $p_1, p_2, \cdots, p_m$. Let $\Omega_\Gamma = \{v \in \Omega |$there exists $A_i \in \Gamma$ such that $v(A_i) = 1, v(p_{m+k}) = 0, k = 1, 2, \cdots\}$, then the element of Ω_Γ is the form as $(v(p_1), v(p_2), \cdots, v(p_m))$ and $|\Omega_\Gamma| \leq 2^m$. For any $v_i \in \Omega_\Gamma, A_j \in \Gamma$,define $v_i I A_j$ if and only if $v_i(A_j) = 1$, then $(\Omega_\Gamma, \Gamma, I)$ forms a formal context, we call the formal context is induced by Γ.

Definition 3. *(Li [5]) Let* $\Gamma \subseteq F(S)$ *be finite, in formal context* $(\Omega_\Gamma, \Gamma, I)$*,* $P \subseteq \Omega_\Gamma, Q \subseteq \Gamma$*. Define the operators* $f(P) = \{A \in \Gamma | vIA$ *for all* $v \in P\}$ *and* $h(Q) = \{v \in \Omega_\Gamma | vIA$ *for all* $A \in Q\}$*.*

From Definition 3, we can prove the following result.

Theorem 2. *(Li [5]) The operators* f *and* h *satisfy a galois connection between* Ω_Γ *and* Γ*, i.e.* $P, P_1, P_2 \subseteq \Omega_\Gamma, Q, Q_1, Q_2 \subseteq \Gamma$*, we have following properties:*

(i) if $P_1 \subseteq P_2$ *then* $f(P_1) \supseteq f(P_2)$*;*
(ii) $P \subseteq hf(P)$*;*
(iii) if $Q_1 \subseteq Q_2$ *then* $h(Q_1) \supseteq h(Q_2)$*;*
(iv) $Q \subseteq fh(Q)$*.*

We immediately deduce the following results by the definition of $(\Omega_\Gamma, \Gamma, I)$.

Theorem 3. *Given a theory* $\Gamma = \{A_1, A_2, \cdots, A_n\} \subset F(S)$*,* Γ *is consistent if and only if there exists* $v \in \Omega_\Gamma$*, such that* $v(\Gamma) = \wedge\{v(A_i) | A_i \in \Gamma\} = 1$*, i.e. in the formal context* $(\Omega_\Gamma, \Gamma, I)$*,* $h(\Gamma) \neq \emptyset$*.*

Theorem 4. *Given a theory* $\Gamma = \{A_1, A_2, \cdots, A_n\} \subset F(S)$*,* $A \in F(S)$*, if* $\Gamma \models A$*, then* $\Gamma \to \{A\}$ *holds in the formal context* $(\Omega_{\Gamma \cup \{A\}}, \Gamma \cup \{A\}, I_{\Gamma \cup \{A\}})$*.*

In order to distinguish two consistent sets Γ and $\sum$ in two-valued propositional logic, the truth degree of a formula and consistency degree are introduced in [9,10].

Definition 4. *(Wang [10,12]) Let $A(p_1, p_2, \cdots, p_n)$ be a formula in L and*

$$\tau(A) = \frac{|\overline{A}^{-1}(1)|}{2^n}.$$

Then $\tau(A)$ is called the truth degree of A, where$\overline{A}^{-1}(1)$ is the total of valuations which satisfies $v(A) = 1$.

Definition 5. *(Wang [9]) Let $\Gamma = \{A_1, A_2, \cdots, A_n\} \subset F(S)$. Define*

$$\delta(\Gamma) = 1 - \tau(A_1 \wedge A_2 \wedge \cdots \wedge A_n \to \overline{0}),$$

then $\delta(\Gamma)$ is called the consistency degree of Γ.

Remark 1. (i) If $\Gamma_0 \subseteq \Gamma$, then $\delta(\Gamma_0) \geq \delta(\Gamma)$. Given a theory $\Gamma = \{A_1, A_2, \cdots, A_n\} \subset F(S)$ contains m-different atomic formulas of S, assume there are $p_1, p_2, \cdots, p_m$. Obviously, for any $A_i \in \Gamma$, $\tau(A_i) \in \{\frac{k}{2^m} | k = 1, 2, \cdots, 2^m\}$ and so is $\delta(\Gamma)$.

(ii) By Definition 4 and Definition 5, we can observe that in if $A \in \Gamma$ and Γ is finite, then $\tau(A) = \frac{|h(\{A\})|}{|\Omega_\Gamma|}$ and $\delta(\Gamma) = \frac{|h(\Gamma)|}{|\Omega_\Gamma|}$ with respect to the formal context $(\Omega_\Gamma, \Gamma, I)$.

Let $\Gamma \subseteq F(S)$. As mentioned above, ε-reduction of theory Γ is to find the minimal theory $\Gamma_0 \subseteq \Gamma$ satisfies that $\delta(\Gamma_0) - \delta(\Gamma) \leq \varepsilon$. we introduce the definition of ε-reduction of theory Γ as follows:

Definition 6. *Let $\Gamma_0 \subseteq \Gamma \subset F(S)$. We say that Γ_0 is a ε-consistent set of theory Γ if $\delta(\Gamma_0) - \delta(\Gamma) \leq \varepsilon$. Moreover, if for any $A \in \Gamma_0$, $\delta(\Gamma_0 - \{A\}) - \delta(\Gamma) > \varepsilon$, then Γ_0 is called a ε-reduction of theory Γ, denote by Γ_ε.*

$\delta(\Gamma) \in \{\frac{k}{2^m} | k = 1, 2, \cdots, 2^m\}$ if Γ contains m-different atomic formulas of S, and $\Gamma_0 \subseteq \Gamma$, so $\delta(\Gamma_0) - \delta(\Gamma) \in \{\frac{k}{2^m} | k = 1, 2, \cdots, 2^m\}$, thus we can only discuss the $\varepsilon \in \{\frac{k}{2^m} | k = 1, 2, \cdots, 2^m\}$.

Definition 7. *Suppose that Γ is a theory, the set $\{\Gamma_{\varepsilon i} | \Gamma_{\varepsilon i}$ is a ε-reduction,$i \in I\}$ (I is an index set) includes all of the ε-reduction of theory Γ, the theory Γ is divided into 3 parts with respect to the error ε:*

(i) necessary formula A: $A \in \bigcap\limits_{i \in I} \Gamma_{\varepsilon i}$;

(ii) useful formula B: $B \in \bigcup\limits_{i \in I} \Gamma_{\varepsilon i} - \bigcap\limits_{i \in I} \Gamma_{\varepsilon i}$;

(iii) useless formula C: $C \in \Gamma - \bigcup\limits_{i \in I} \Gamma_{\varepsilon i}$.

Theorem 5. *The ε-reduction exists for any finite theory Γ at a given error ε.*

Proof. Let Γ be a finite theory. If for any $A \in \Gamma$, $\delta(\Gamma - \{A\}) - \delta(\Gamma) > 0$, then Γ itself is a 0-reduction. If there is a formula $A \in \Gamma$ such that $\delta(\Gamma - \{A\}) - \delta(\Gamma) = 0$, then we study $\Gamma_0 = \Gamma - \{A\}$, Further, if for any $B \in \Gamma_0$, $\delta(\Gamma_0 - \{B\}) - \delta(\Gamma_0) > 0$, then Γ_0 is a 0-reduction, otherwise we study $\Gamma_0 - \{B\}$. Repeating the above process, we can find one 0-reduction because Γ is finite. So 0-reduction of Γ must exist. Thus ε-reduction exists by Definition 6.

3 Judgment of ε-Reduction of Theory Γ and Some Properties

As mentioned above, $\Gamma_0 \subseteq \Gamma$ is a ε-reduction if and only if: Γ_0 is a ε-consistent set and $\Gamma_0 \setminus \{A\}$ is not a ε-consistent set for any $A \in \Gamma_0$. So we only need to give the judgment theory of ε-consistent set.

Theorem 6. *$\Gamma_0 \subseteq \Gamma$, Γ_0 is a ε-consistent set if and only if $\Gamma_0 \to \Gamma \setminus \Gamma_0$ is hold in context $(\Omega_\Gamma, \Gamma, I)$ with respect to the confidence threshold $\frac{\delta(\Gamma)}{\delta(\Gamma)+\varepsilon}$.*

Proof. Necessity. Because Γ_0 is a ε-consistent set, it is easy to see that $\delta(\Gamma_0) \leq \delta(\Gamma)+\varepsilon$, that is to say, $\frac{|h(\Gamma_0)|}{|\Omega_\Gamma|} \leq \delta(\Gamma)+\varepsilon$. Since $\Gamma_0 \subseteq \Gamma$, $h(\Gamma_0) \supseteq h(\Gamma)$ by Theorem 2, thus $\frac{|h(\Gamma_0 \cup \Gamma \setminus \Gamma_0)|}{|h(\Gamma_0)|} = \frac{|h(\Gamma)|}{|h(\Gamma_0)|} = \frac{\frac{|h(\Gamma)|}{|\Omega_\Gamma|}}{\frac{|h(\Gamma_0)|}{|\Omega_\Gamma|}} = \frac{\delta(\Gamma)}{\delta(\Gamma_0)} \geq \frac{\delta(\Gamma)}{\delta(\Gamma)+\varepsilon}$. We conclude that $\Gamma_0 \to \Gamma \setminus \Gamma_0$ is hold in context $(\Omega_\Gamma, \Gamma, I)$ with respect to the confidence threshold $\frac{\delta(\Gamma)}{\delta(\Gamma)+\varepsilon}$.

Sufficiency. Suppose that $\Gamma_0 \to \Gamma \setminus \Gamma_0$ is hold in context $(\Omega_\Gamma, \Gamma, I)$ with respect to the confidence threshold $\frac{\delta(\Gamma)}{\delta(\Gamma)+\varepsilon}$, $\frac{|h(\Gamma_0 \cup \Gamma \setminus \Gamma_0)|}{|h(\Gamma_0)|} \geq \frac{\delta(\Gamma)}{\delta(\Gamma)+\varepsilon}$ by Definition 2, i.e. $\frac{|h(\Gamma)|}{|h(\Gamma_0)|} = \frac{\frac{|h(\Gamma)|}{|\Omega_\Gamma|}}{\frac{|h(\Gamma_0)|}{|\Omega_\Gamma|}} = \frac{\delta(\Gamma)}{\delta(\Gamma_0)} \geq \frac{\delta(\Gamma)}{\delta(\Gamma)+\varepsilon}$, so $\delta(\Gamma_0) \leq \delta(\Gamma) + \varepsilon$, We conclude that Γ_0 is a ε-consistent set by Definition 6.

Theorem 7. *Let $\Gamma_0 \subseteq \Gamma$ and Γ_0 is a ε-consistent set of Γ. Γ_0 is a ε-reduction if and only if for any $A \in \Gamma_0, \Gamma_0 \setminus \{A\} \to \{A\}$ is not hold with respect to the confidence threshold $\varphi = 1$.*

Proof. Γ_0 is a ε-reduction, if and only if for any $A \in \Gamma_0$, $\Gamma_0 \setminus \{A\}$ is not a ε-consistent set, if and only if $\delta(\Gamma_0 \setminus \{A\}) - \delta(\Gamma) > \varepsilon$, i.e. $\delta(\Gamma_0 \setminus \{A\}) - \delta(\Gamma_0) + \delta(\Gamma_0) - \delta(\Gamma) > \varepsilon$, since Γ_0 is a ε-consistent set, $\delta(\Gamma_0) - \delta(\Gamma) \leq \varepsilon$, then $\delta(\Gamma_0 \setminus \{A\}) - \delta(\Gamma_0) > 0$, if and only if $\frac{|h(\Gamma_0 \setminus \{A\})|}{\Omega_\Gamma} - \frac{|h(\Gamma_0|}{\Omega_\Gamma} > 0$, if and only if $\frac{|h(\Gamma_0)|}{|h(\Gamma_0 \setminus \{A\})|} < 1$, we have complete the proof.

Theorem 8. *Let $\Gamma_1 \subseteq \Gamma_0 \subseteq \Gamma$ and $\varepsilon_1 \leq \varepsilon_2$, Γ_0 is a ε_1-reduction of Γ. Then Γ_1 is a ε_2-consistent set of Γ if and only if Γ_1 is a ε_2-reduction of Γ.*

Proof. The sufficiency is obviously, we have to prove the necessity. Suppose that Γ_1 is a ε_2-consistent set of Γ, for any $A \in \Gamma_1 \subseteq \Gamma_0$, $\Gamma_0 \setminus \{A\} \to \{A\}$ is not hold with respect to the confidence threshold $\varphi = 1$ by Γ_0 is a ε_1-reduction of Γ. Since $\Gamma_1 \subseteq \Gamma_0$ and $h(\Gamma_1) \supseteq h(\Gamma_0)$, $h(\Gamma_1 \setminus \{A\}) \supseteq h(\Gamma_0 \setminus \{A\})$. Suppose that $\Gamma_1 \setminus \{A\} \to \{A\}$ is hold with respect to the confidence threshold $\varphi = 1$, then $h(\Gamma_1 \setminus \{A\}) \subseteq h(\{A\})$. So $h(\Gamma_0 \setminus \{A\}) \subseteq h(\Gamma_1 \setminus \{A\}) \subseteq h(\{A\})$, thus $\Gamma_0 \setminus \{A\} \to \{A\}$ is hold with respect to the confidence threshold $\varphi = 1$, it is a contradiction and we have proved the conclusion.

Theorem 9. *Let $\Gamma_1 \subseteq \Gamma_0 \subseteq \Gamma$, Γ_0 is a ε_1-reduction of Γ, Γ_1 is a ε_2-reduction of Γ_0. Then Γ_1 is a $\varepsilon_1 + \varepsilon_2$-reduction of Γ.*

Proof. Since Γ_0 is a ε_1-reduction of Γ, then $\delta(\Gamma_0) - \delta(\Gamma) \leq \varepsilon_1$, that is $\frac{|h(\Gamma_0)|-|h(\Gamma)|}{|\Omega_\Gamma|} \leq \varepsilon_1$. $\frac{|h(\Gamma_1)|-|h(\Gamma_0)|}{|\Omega_\Gamma|} \leq \varepsilon_2$ by the similarly way. Thus $\frac{|h(\Gamma_1)|-|h(\Gamma)|}{|\Omega_\Gamma|} \leq \varepsilon_1 + \varepsilon_2$, that is to say $\delta(\Gamma_1) - \delta(\Gamma) \leq \varepsilon_1 + \varepsilon_2$, so Γ_1 is a $\varepsilon_1 + \varepsilon_2$-reduction of Γ.

Suppose that Γ is a theory, the set $\{\Gamma_{\varepsilon i} | \Gamma_{\varepsilon i}$ is a ε-reduction$i \in I\}$(I is an index set) includes all of the ε-reduction of theory Γ. From Definition 7 and Theorem 6, The following theorem is easily obtained and the proof is omitted.

Theorem 10. *$A \in \bigcap_{i \in I} \Gamma_{\varepsilon i}$ is a necessary formula if and only if $\Gamma \setminus \{A\} \to \{A\}$ is not hold with respect to the confidence threshold $\varphi = \frac{\delta(\Gamma)}{\delta(\Gamma)+\varepsilon}$.*

Corollary 1. *$\bigcap_{i \in I} \Gamma_{\varepsilon i}$ is a ε-reduction if and only if there if only one ε-reduction.*

Corollary 2. *Suppose $A, B \in \Gamma$, if $A \in \bigcap_{i \in I} \Gamma_{\varepsilon i}$ and the confidence threshold of $\{A\} \to \{B\}$ is $\varphi = 1$,then B is an useless formula.*

4 Methods to Explore the ε-Reduction

In this section, two methods of exploring ε-reduction will be given based on the Theorems 6-10 in Section 3.

From Theorem 6 and Theorem 7, in order to explore the ε-reduction of Γ, we can obtain a method as follows:

(i) Given a theory $\Gamma = \{A_1, A_2, \cdots, A_n\} \subset F(S)$ contains m-different atomic formulas of S, calculate the $(\Omega_\Gamma, \Gamma, I)$ by rules of two-valued logic.

We only need to explore the association rules at the given confidence threshold $\varphi = \frac{\delta(\Gamma)}{\delta(\Gamma)+\varepsilon}$ that satisfy the following rules:

(ii) Find the subset Γ_0 satisfy $\Gamma_0 \to \Gamma \setminus \Gamma_0$ is hold with respect to the confidence threshold $\varphi = \frac{\delta(\Gamma)}{\delta(\Gamma)+\varepsilon}$.

(iii) For any $A \in \Gamma_0, \Gamma_0 \setminus \{A\} \to \{A\}$ is not hold with respect to the confidence threshold $\varphi = 1$.

Table 1. Formal context $(\Omega_\Gamma, \Gamma, I)$ of Example 1

I	A_1	A_2	A_3	A_4	A_5	A_6
v(0,0,0)	0	1	0	1	0	1
v(1,0,0)	1	1	1	1	0	1
v(0,1,0)	0	0	0	0	1	0
v(0,0,1)	0	1	1	1	1	1
v(1,1,0)	1	0	0	0	1	1
v(1,0,1)	1	1	1	1	1	1
v(0,1,1)	0	1	1	0	0	0
v(1,1,1)	1	1	1	0	0	0

Example 1. Let $\Gamma = \{p_1, p_1 \to p_2, (p_1 \to p_2) \to p_3, \neg p_2, (\neg p_2 \to p_3) \wedge (p_3 \to \neg p_2), p_2 \to p_1 \wedge p_3\} = \{A_1, A_2, \cdots, A_6\}$. Then formal context $(\Omega_\Gamma, \Gamma, I)$ is shown in table 1, and $\delta(\Gamma) = \frac{1}{8}$ by Definition 5.

If $\varepsilon = 0$(confidence threshold$\varphi = 1$), the association rules satisfies (ii) and (iii) as follows:

$\{A_1, A_2, A_5\} \to \{A_3, A_4, A_6\}$,
$\{A_1, A_3, A_5\} \to \{A_2, A_4, A_6\}$,
$\{A_1, A_4, A_5\} \to \{A_2, A_3, A_6\}$,

thus, 0-reduction of Γ are $\{A_1, A_2, A_5\}$,$\{A_1, A_3, A_5\}$ and $\{A_1, A_4, A_5\}$, on this condition A_1, A_5 are necessary formulas , A_2, A_3, A_4 are useful formulas and A_6 is useless formula.

If $\varepsilon = \frac{1}{8}$(confidence threshold$\varphi = \frac{1}{2}$), the association rules satisfies (ii) and (iii) as follows:

$\{A_1, A_5\} \to \{A_2, A_3, A_4, A_6\}$,
$\{A_3, A_5\} \to \{A_1, A_2, A_4, A_6\}$,
$\{A_1, A_4\} \to \{A_2, A_3, A_5, A_6\}$,
$\{A_2, A_5\} \to \{A_1, A_3, A_4, A_6\}$,
$\{A_4, A_5\} \to \{A_1, A_2, A_3, A_6\}$,

thus, $\frac{1}{8}$-reduction of Γ are $\{A_1, A_5\}$,$\{A_3, A_5\}$, $\{A_1, A_4\}$, $\{A_2, A_5\}$ and $\{A_4, A_5\}$,on this condition there is no necessary formula , A_1, A_2, A_3, A_4, A_5 are useful formulas and A_6 is useless formula.

If $\varepsilon = \frac{2}{8}$(confidence threshold$\varphi = \frac{1}{3}$), the association rules satisfies (ii) and (iii) as follows:

$\{A_1, A_2\} \to \{A_3, A_4, A_5, A_6\}$,
$\{A_1, A_3\} \to \{A_2, A_4, A_5, A_6\}$,
$\{A_1, A_6\} \to \{A_2, A_3, A_4, A_5\}$,
$\{A_3, A_4\} \to \{A_1, A_2, A_5, A_6\}$,
$\{A_3, A_6\} \to \{A_1, A_2, A_4, A_5\}$,

thus, $\frac{2}{8}$-reduction of Γ are $\{A_1, A_2\}$,$\{A_1, A_3\}$, $\{A_1, A_6\}$, $\{A_3, A_4\}$ and $\{A_3, A_6\}$,on this condition there is no necessary formula , A_1, A_2, A_3, A_4, A_6 are useful formulas and A_5 is useless formula.

By similarly way, we can find ε-reductions when ε even larger.

From Theorem 9 and Theorem 10, we can obtain an other way to explore a ε-reduction of Γ as follows:

(i) Given a theory $\Gamma = \{A_1, A_2, \cdots, A_n\} \subset F(S)$ contains m-different atomic formulas of S, calculate the $(\Omega_\Gamma, \Gamma, I)$ by rules of two valued logic.

(ii) Let $\Gamma_0 = \Gamma$. Suppose the confidence threshold of $\Gamma_0 \setminus \{A_1\} \to \{A_1\}$ is θ_1, if $\frac{\delta(\Gamma_0)}{\theta_1} - \delta(\Gamma_0) \geq \varepsilon$, that is to say $A_1 \in \Gamma_0$ is not a necessary formula of Γ_0 with the error ε, then let $\Gamma_1 = \Gamma_0 \setminus \{A_1\}$, else $\Gamma_1 = \Gamma_0$. For $A_2 \in \Gamma_1$ and $A_2 \neq A_1$, suppose the confidence threshold of $\Gamma_1 \setminus \{A_2\} \to \{A_2\}$ is θ_2, if $\frac{\delta(\Gamma_0)}{\theta_1} - \delta(\Gamma_0) + \frac{\delta(\Gamma_1)}{\theta_2} - \delta(\Gamma_1) \geq \varepsilon$, that is to say $A_2 \in \Gamma_1$ is not a necessary formula of Γ_1 with the error $\varepsilon - \frac{\delta(\Gamma_0)}{\theta_1} + \delta(\Gamma)$, then let $\Gamma_2 = \Gamma_1 \setminus \{A_2\}$, else $\Gamma_2 = \Gamma_1$. And so on, we can get a theory Γ_i satisfies that every formula in Γ_i is necessary formula of Γ_i with the error of $\varepsilon - \frac{\delta(\Gamma_0)}{\theta_1} - \frac{\delta(\Gamma_1)}{\theta_2} - \cdots - \frac{\delta(\Gamma_{i-1})}{\theta_i} + \delta(\Gamma_1) + \delta(\Gamma_2) + \cdots + \delta(\Gamma_{i-1})$ since Γ is finite. The Γ_i is exactly a ε-reduction of Γ.

Example 2. Let us take the formal context $(\Omega_\Gamma, \Gamma, I)$ described as Table 1 as an example.

If $\varepsilon = \frac{1}{8}$, then confidence threshold $\varphi = \frac{1}{2}$. $\Gamma_0 = \Gamma$, A_1 is not a necessary formula of Γ_0 with the error $\frac{1}{8}$ since the confidence threshold of $\{A_2, A_3, A_4, A_5, A_6\} \to \{A_1\}$ is $\frac{1}{2}$ and $\frac{\delta(\Gamma_0)}{\theta_1} - \delta(\Gamma_0) = \frac{\frac{1}{8}}{\frac{1}{2}} - \frac{1}{8} = \frac{1}{8} = \varepsilon$, then $\Gamma_1 = \Gamma_0 \setminus \{A_1\}$. Since $\varepsilon - \frac{\delta(\Gamma_0)}{\theta_1} + \delta(\Gamma_0) = 0$, we have to eliminate the unnecessary formulas in Γ_1 with the error 0 step by step. Since the confidence threshold of $\{A_3, A_4, A_5, A_6\} \to \{A_2\}$ is 1, then $\Gamma_2 = \Gamma_1 \setminus \{A_2\}$. In a similar vein, A_3 is not a necessary formula of Γ_2 with the error of 0, so $\Gamma_3 = \Gamma_2 \setminus \{A_3\}$. Since the confidence threshold of $\{A_5, A_6\} \to \{A_4\}$ is $\frac{2}{3}$ and so $\{A_4\}$ is a necessary formula of Γ_3 with the error 0, so $\Gamma_4 = \Gamma_3$. The confidence threshold of $\{A_4, A_6\} \to \{A_5\}$ is $\frac{1}{3}$ and so $\{A_5\}$ is a necessary formula of Γ_4 with the error 0, so $\Gamma_5 = \Gamma_4$. A_6 is not a necessary formula of Γ_5 with the error 0. Thus $\Gamma_6 = \Gamma_5 \setminus \{A_6\} = \{A_4, A_5\}$ is a $\frac{1}{8}$−reduction of Γ.

Remark 2. From the context, we know that whether a formula is an necessary or not depends on the error ε, but the transitivity dose not hold. That is, a necessary formula with respect to a large error is not necessarily a necessary formula with respect to a small error and similarly, an useful formula with respect to a large error is not necessarily an useful formula with respect to a small error and so is useless formula. We can observe the fact in Example 1.

5 Conclusion

In this paper, the mutual relationship between association rules and theory reduction are investigated by introducing the formal context $(\Omega_\Gamma, \Gamma, I)$, then the theory of ε-reduction of the finite theory Γ in two valued propositional logic is proposed, and we divided the theory Γ into 3 parts with respect to the error ε: necessary formula, useful formula and useless formula. By introducing the association rules to the formal context $(\Omega_\Gamma, \Gamma, I)$, judgment theorems of ε-consistent theorems are examined, and two approaches to explore ε-reduction are presented. In the present paper, it seems that just the confidence threshold is considered, the fact is that the support threshold is determined at the given theory Γ,i.e. support threshold of $\Gamma_0 \to \Gamma \setminus \Gamma_0$ is $\theta = \frac{|h(\Gamma)|}{\Omega_\Gamma} = \delta(\Gamma)$. The ε-reduction of n-valued systems and continuous-valued system is an interesting and important problems and we consider those problems in other papers.

Acknowledgements. This work is supported by the National Natural Science Foundation of China (Grant No. 60803162), and the Science Plan Foundation of the Education Bureau of Shaanxi Province (No.09JK722;09JK439), and Shaanxi Provincial Science and Technology Research and Development Projects (No. 2008K01-33).

References

1. Liberatore, P.: Redundancy in logic I: CNF propositional formulae. Artificition Intelligence 163(2), 203–232 (2005)
2. Liberatore, P.: Redundancy in logic II: 2-CNF and Horn propositional formulae. Artificition Intelligence 172(2-3), 265–299 (2008)
3. Liberatore, P.: Redundancy in logic III: Non-monotonic reasoning. Artificition Intelligence 172(11), 1317–1359 (2008)
4. Ren, Y., Ma, X.J., Wang, H.T.: Reduction of a propositional set and its roots in $\mathcal{L}^*$ system. Fuzzy Systems and Mathematics 20(2), 23–27 (2006) (in Chinese)
5. Li, L.F., Zhang, D.X.: The application of concept lattice theory in the reduction of the proposition set in two-valued propositional logic. Acta Electronica Sinica 35(8), 1538–1542 (2007) (in Chinese)
6. Agrawl, R., Imielinski, A.: Mining association rules between sets of items in large databases. In: Proc. ACM SIGMOD International Conference on Management of Data, Washington, D.C., May 1993, pp. 207–216 (1993)
7. Pasquier, N., Taouil, R., Bastide, Y., Stumme, G., Lakhal, L.: Generating a condensed representation for association rules. Journal of Intelligent Information Systems 24(1), 29–60 (2005)
8. Zaki, M.J., Hsiao, C.-J.: Efficient Algorithms for Mining Closed. Itemsets and Their Lattice Structure. IEEE Transactions on Knowledge and Data Engineering 17(4), 462–478 (2005)
9. Zhou, X.N., Wang, G.J.: Consistency degrees of theories in some systems of propositional fuzzy logic. Fuzzy Sets and System 152(3), 321–331 (2005)
10. Wang, G.J.: Introduction to mathematical logic and resolution principle. Science in China Press, Beijing (2006) (in Chinese)
11. Ganter, B., Wille, R.: Formal Concept Analysis: Mathematical Foundations. Springer, Berlin (1999)
12. Wang, G.J., Fu, L., Song, J.S.: Theory of truth degrees of propositions in two-valued logic. Science in China, Ser. A 45(9), 1106–1116 (2002)

A Family $I_{L\Pi GN(q,p)}$ of Implication Operators and Triple I Methods Based on It

Xing-fang Zhang

School of Mathematical Sciences, Liaocheng University, Liaocheng 252059, China
e-mail: zhangxingfang2005@126.com

Abstract. In this paper, first, a new family $T_{L\Pi GN(q,p)}, q \in [-1,1], P \in (-\infty,+\infty)$ of left -continuous t-norms are presented; and then its residual family $I_{L\Pi GN(q,p)}, q \in [-1,1], P \in (-\infty,+\infty)$ of implication operators are given; finally, a generic form of Triple I methods based on the family $I_{L\Pi GN(q,p)}$ of implication operators in fuzzy reasoning is expressed.

Keywords: Fuzzy reasoning, left-continuous t-norm, implication operator, Triple I method.

1 Introduction

Triangular norms (briefly t-norms) are an indispensable tool in the interpretations of the conjunction in fuzzy logics [1] and the intersection of fuzzy sets [2,3]. They also play an important role in decision making [4], statistics analysis [5,6], the theories of non-additive measures [7,8], cooperative games [9] and the solutions of well-known functional equations [10]. In particular, in CRI method of fuzzy reasoning presented by Zádeh [2,3] and Triple I method of fuzzy reasoning presented by Wang [11,12,13], reasoning results nearly relate to t-norms and implication operators.

It is well known that fuzzy modus ponens (briefly, FMP) and fuzzy modus tolens (briefly, FMT) can be

for a given rule $A \to B$ and input A^* calculate B^*, {1}

for a given rule $A \to B$ and input B^* calculate A^*, {2}

respectively, where $A, A^* \in F(X)$ (the set of all fuzzy subsets on universe X) and $B, B^* \in F(Y)$ (the set of al l fuzzy subsets on universe Y). At present, the most widespread mean of solving such problem is Zadeh's CRI (Compositional Rule of Inference) method proposed in 1973 [3]. The method on FMP translates $A(x) \to B(y)$ into a fuzzy binary relation $R(A(x), B(y))$ and combines A^* and $R(A(x), B(y))$, that is:

$$B^*(y) = \sup_{x \in X} \{A^*(x) * R(A(x), B(y))\}, y \in Y, \quad (1)$$

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 167–175.
springerlink.com

where $*$ is a composite operation. Based on point of view of logic Wang [3] has proposed the triple I (the abbreviation of triple implications) method of solving the FMP (FMT) problems as follows:

Definition 1. *([11,12]) Let* X,Y *be nonempty sets.* $F(X), F(Y)$ *denote the family of all fuzzy subsets of* X *and* Y *respectively. Given fuzzy sets* $A(x) \in F(X), B(y) \in F(y)$, $A^*(x) \in F(X)(B^*(y) \in F(Y))$ *Then the methods of seeking the minimum fuzzy set* $B^*(y) \in F(Y)(A^*(x) \in F(X)$, *such that*

$$(A(x) \Rightarrow B(y)) \Rightarrow (A^*(x) \Rightarrow B^*(y))$$

has maximum(minimum)possible value, for any $x \in X$ *and* $y \in Y$, *is said to be Triple I (Triple I for short) method for FMP (FMT), where* $\Rightarrow$ *is an implication and the* $B^*(y)(A^*(x))$ *is called the solution of Triple I for FMP (FMT). If* $\Rightarrow$ *is the residual implication generated by a left-continuous t-norm, then solution of Triple I for FMP (FMT) is given by the following formulas.*

Theorem 1. *([11-13]) Let the operator* I *in FMP* $\{1\}$ *is the residual implication* I *generated by a left-continuous t-norm* $*$. *Then the triple I solution on FMP is*

$$B^*(y) = \sup_{x \in X} \{A^*(x) * I(A(x), B(y))\}, y \in Y \tag{2}$$

Theorem 2. *([11-13]) Let the operator* I *in FMT* $\{2\}$ *is the residual implication* I *generated by a left-continuous t-norm* $*$. *Then the triple I solution on FMT is*

$$A^*(y) = \inf_{x \in X} \{I(I(A(x), B(y),\ B^*(y))\}, x \in X \tag{3}$$

In CRI method or Triple I method of fuzzy reasoning, reasoning result based on different implications are popularly different. Their difference are often very great. Therefore, people use implication operators with parameter (or family of implication operators) in fuzzy reasoning to reduce the reasoning error. These papers [5-8] have given some family of implication operators. The paper again proposes a new family of t-norms denoted by $T_{L\Pi GN(q,p)}, q \in [-1, 1], P \in (-\infty, +\infty)$ and its residual the family of implications operators denoted by $I_{L\Pi GN(q,p)}, q \in [-1, 1], P \in (-\infty, +\infty)$. It contains all excellent implication operators: Lukasiewicz implication operator, Godel implication operator, NM implication operator and product implication operator. Finally, a generic form of Triple I methods based on the family $I_{L\Pi GN(q,p)}$ of implication operators is expressed.

2 Preliminaries

Definition 2. *([1]) A t-norm* $*$ *is a binary operation on* $[0, 1]$ *(i.e.,* $* : [0, 1]^2 \to [0, 1]$*) satisfying the following conditions* :

(i) $*$ *is commutative, i.e., for all* $x, y \in [0,1]$, $(x * y) = y * x$;
(ii) $*$ *is associative, i.e., for all* $x, y, z \in [0,1]$, $(x * y) * z = x * (y * z)$;
(iii) $*$ *is non-decreasing in both arguments, i.e., for all* $x, y, z \in [0,1]$,
$y \leq z$ *implies* $x * y \leq x * z$, $y \leq z$ *implies* $y * x \leq z * x$;
(iv) $x * 1 = x$, *for all* $x \in [0,1]$.

The above t-norm is also denoted by $T(x,y), (x,y) \in [0,1]$.

A binary operation $\Rightarrow$ on $[0,1]$ is called the residual implication of t-norm $*$, if $x * y \leq z$ if and only if $x \leq y \Rightarrow z$, for all $x, y, z \in [0,1]$. The residual implication of t-norm $*$ is denoted by $\Rightarrow_*$.

Lukasiewicz t-norm $*_{Lu}$ and its residual implication $\Rightarrow_{Lu}$:

$$x *_{Lu} y = (x + y - 1) \vee 0, x, y, \in [0,1], \tag{4}$$

$$x \Rightarrow_{Lu} y = (1 - x + y) \wedge 1, x, y, \in [0,1]. \tag{5}$$

Gödel t-norm $*_G$ and its residual implication $\Rightarrow_G$:

$$x *_G y = x \wedge y, x, y, \in [0,1], \tag{6}$$

$$x \Rightarrow_G y = \begin{cases} 1, \text{ if } x \leq y, x, y \in [0,1], \\ y, \text{ if } x > y, x, y \in [0,1]. \end{cases} \tag{7}$$

Product t-norm $*_\Pi$ and its residual implication $\Rightarrow_\Pi$:

$$x *_\Pi y = x \times y, x, y, \in [0,1], \tag{8}$$

$$x \Rightarrow_\Pi y = \begin{cases} 1, \text{ if } x \leq y, x, y \in [0,1], \\ \dfrac{y}{x}, \text{ if } x > y, x, y \in [0,1]. \end{cases} \tag{9}$$

NM t-norm $*_{NM}$ and its residual implication $\Rightarrow_{NM}$:

$$x *_{NM} y = \begin{cases} x \wedge y, \text{ if } x + y > 1, x, y \in [0,1], \\ 0, \quad \text{if } x + y \leq 1, x, y \in [0,1], \end{cases} \tag{10}$$

$$x \Rightarrow_{NM} y = \begin{cases} 1, & \text{if } x \leq y, x, y \in [0,1], \\ (1 - x) \vee y, & \text{if } x > y, x, y \in [0,1]. \end{cases} \tag{11}$$

3 Family $T_{L\Pi GN(q,p)}$ of t-Norms and Its Residual Family $I_{L\Pi GN(q,p)}$ of Implication Operators

Theorem 3. *For any* $q \in [-1,1]$, $P \in (-\infty, +\infty)$, *the binary operation* $*_{(q,p)}$ *satisfying*

$$x *_{(q,p)} y = \begin{cases} \lim\limits_{t \to p} (x^t + y^t - q)^{\frac{1}{t}} \cap x \cap y, x^p + y^p > |q| \textit{ and } xy \neq 0, \\ 0 \qquad\qquad , \textit{otherwise}, \end{cases}$$

$(x,y) \in [0,1] \times [0,1]$ *are left- continuous t-norms.*

Proof. We only prove the cases of $q \in [0,1], P \in (0,\infty)$ since similarly as it we can prove other cases.

According to the definition of t-norm, we should verify that, for any $q \in [0,1], P \in (0,\infty)$, the binary operation $*_{(q,p)}$ (simple written as $*$ in the following) hold the following properties.

(i) $x * y = y * x$, for all $(x,y) \in [0,1] \times [0,1]$;

(ii) $*$ is non-decreasing for all $(x,y) \in [0,1] \times [0,1]$;

(iii)$x * 1 = x$, for all $x \in [0,1]$;

(iv) $(x * y) * z = x * (y * z)$, for all x, y, z.

It is easy to verify the properties (i)-(iii). In the following we only prove the property (iv), i.e., the binary operation $*$ is associative.

1) Assume $x^p + y^p > q, x^p + z^p > q, y^p + z^p > q$.

If $x^p > q, y^p > q,\ z^p > q$, then

$$(x * y) * z = x \cap y \cap z, x * (y * z) = x \cap y \cap z.$$

If $x^p \leq q, y^p > q, z^p > q$, then

$$(x * y) * z = x * z = x,\ x * (y * z) = x * (y \cap z) = x \cap (y \cap z) = x.$$

If $x^p \leq q, y^p \leq q, z^p > q$, then

$$(x * y) * z = (x^p + y^p - q)^{\frac{1}{p}} * z = (x^p + y^p - q)^{\frac{1}{p}} \cap z = (x^p + y^p - q)^{\frac{1}{p}},$$

$$x * (y * z) = x * y = (x^p + y^p - q)^{\frac{1}{p}}.$$

If $x^p \leq q, y^p > q, z^p \leq q$, then

$$(x * y) * z = a * c = (x^p + c^p - q)^{\frac{1}{p}},$$

$$x * (y * z) = x * z = (x^p + z^p - q)^{\frac{1}{p}}.$$

If $x^p \leq q, y^p \leq q, z^p \leq q$, then

$$x * y * z = (x^p + y^p - q)^{\frac{1}{p}} * z = (x^p + y^p + z^p - 2q)^{\frac{1}{p}},$$

$$x * (y * z) = x * (y^p + z^p - q)^{\frac{1}{p}} = (x^p + y^p + z^p - 2q)^{\frac{1}{p}}.$$

2) Assume $x^p + y^p \leq q$. Then $x * y * z = 0 * z = 0$, $x * (y * z) \leq x * (b * 1) = x * y = 0$. So

$$x * y * z = x * (y * z) = 0.$$

Similarly, we can prove the cases of $x^p + z^p \leq q$ and $z^p + z^p \leq q$.

Therefore these binary operations $*$ are all t-norms.

Next, we prove that for any $q \in [0,1]$, $P \in (0,\infty)$, the operators $*$ are left-continuous, i.e., $\forall x \in [0,1]$, $x * (\cup E) = \cup(x * E)$, where $E = \{e|0 \leq e < d\}$, $\cup E = Sup\{e|0 \leq e < d\} = d$, $(\cup E^p) = Sup\{e^p|0 \leq e < d\} = d^p$.

In fact, 1) When $x^p + d^p \leq q$, it follows from $x * (\cup E) = 0$ that $x^p + e^p \leq q$, for any $e \in E$. Hence $x * e = 0$. Thus $\cup(x * E) = 0$.

2) When $x^p + d^p > q$, it follows from $*$ and $\cap$ are all monotone that

$$x * d = (x^p + d^p - q)^{\frac{1}{p}} \cap x \cap d = (x^p + (\cup E)^p - q)^{\frac{1}{p}} \cap (x \cap (\cup E)),$$
$$\cup(x^p + E^p - q)^{\frac{1}{p}} \cap [\cup(x \cap E)] = \cup[(x^p + E^p - q)^{\frac{1}{p}} \cap (x \cap E)]$$
$$= \cup[(x^p + E^p - q)^{\frac{1}{p}} \cap x \cap E] = \cup[(x^p + E^p - q)^{\frac{1}{p}} \cap x \cap E] = \cup(x * E),$$

where $x \cap E$ and $x * E$ are the shortening of $\{x \cap e | e \in E\}$ and $\{x * e | e \in E\}$, respectively.

Therefore, for any $q \in [0,1]$, $P \in (0, \infty)$, $*_{(q,p)}$ are left-continuous t-norms.

In conclusion, for any $q \in [-1,1]$, $p \in (-\infty, +\infty)$, $*_{(q,p)}$ (i.e., $T_{(q,p)}(x,y)$) are left -continuous t-norms.

Note that when $(q,p) = (1,1)$, $(q,p) = (0,1)$, $(q,p) = (-1,1)$, $(q,p) = (1,0)$,

$$T_{(q,p)}(x,y) = \begin{cases} \lim\limits_{t \to p} (x^t + y^t - q)^{\frac{1}{t}} \cap x \cap y, x^p + y^p > |q| \ and \ xy \neq 0 \\ \qquad 0, \qquad \text{otherwise}, \end{cases}$$

$(x,y) \in [0,1] \times [0,1]$ are t-norms T_L , T_G, T_{NM} and T_Π, respectively. Thus, t-norms in Theorem 3 are denoted by $*_{L\Pi GN(q,p)}(x,y), (x,y) \in [0,1]$ or $T_{L\Pi GN(q,p)}(x,y), (x,y) \in [0,1]$, and the following definition is given.

Definition 3. *The set* $\{T_{L\Pi GN(q,p)}(x,y), (x,y) \in [0,1] | (q,p) \in [-1,1] \times (-\infty, +\infty)\}$ *is called the family of t-norms. We use symbol* $T_{L\Pi GN(q,p)}(x,y)$, $(x,y) \in [0,1], (q,p) \in [-1,1] \times (-\infty, +\infty)$, *or briefly* $T_{L\Pi GN(q,p)}$, *to denote it.*

Note that the class of t-norm

$$T_{L\Pi GN(q,p)}(x,y) = \begin{cases} x \wedge y, x^p + y^p > |q| \\ 0, x^p + y^p \leq |q| \end{cases}, (x,y) \in [0,1] \times [0,1],$$

$(q,p) \in [-1,0] \times (0, \infty)$ is denoted by $T_{GN(q,p)}$;
the class of t-norm

$$T_{(q,p)}(x,y) = \begin{cases} \left.\begin{array}{c} x \wedge y, x^p > q \ \ or y^p > q \\ (x^p + y^p - q)^{\frac{1}{p}}, x^p \leq q \ \ and y^p \leq q \end{array}\right\} x^p + y^p > q, \\ \qquad 0, x^p + y^p \leq q, \end{cases}$$

$(x,y) \in [0,1] \times [0,1], (q,p) \in [0,1] \times (0, \infty)$, is denoted by $T_{LG(q,p)}$;

when $(q,p) \in [-1,1] \times (-\infty, 0)$ $T_{(q,p)}(x,y) = \begin{cases} (x^p + y^p - q)^{\frac{1}{p}}, xy \neq 0 \\ 0, xy = 0, \end{cases}$

when $q \in [-1,1), p = 0$, $x *_{(q,p)} y = x \wedge y$; when $q = 1, p = 0$, $x *_{(q,p)} y = xy$.

Theorem 4. *The residual implications* $\Rightarrow_{(q,p)}$ *of* $T_{L\Pi GN(q,p)}, (q,p) \in [-1,1] \times (-\infty, +\infty)$ *is given by*

$$x \Rightarrow_{(q,p)} y = \begin{cases} 1, x \leq y \\ (1-f+fy) \vee \lim\limits_{t \to p} ((|q|-x^t) \vee 0)^{\frac{1}{|t|}} f, q - x^p + y^p \leq 0, x > y \\ \lim\limits_{t \to p^+} (q - x^t + y^t)^{\frac{1}{t}} \vee fy \vee \lim\limits_{t \to p} ((|q|-x^t) \vee 0)^{\frac{1}{|t|}} f, q - x^p + y^p > 0, x > y, \end{cases}$$

where $f = \begin{cases} 1, p \geq 0 \\ 0, p < 0 \end{cases}$ *, stipulate that* $t = 0, 0^t = 0^{\frac{1}{t}} = 0.$

Note that for $T_{L\Pi GN(q,p)}$, when $(q,p) \in [-1,0] \times (0,\infty)$, we have

$$T_{(q,p)}(x,y) = \begin{cases} 1, x \leq y \\ y \vee ((|q| - x^p) \vee 0)^{\frac{1}{p}}, x > y \end{cases}, (x,y) \in [0,1] \times [0,1];$$

when $(q,p) \in [0,1] \times (0,\infty)$, we have

$$T_{(q,p)}(x,y) = \begin{cases} 1, x \leq y \\ \left. \begin{array}{l} y, x^p > q \\ (q - x^p + y^p)^{\frac{1}{p}}, x^p \leq q \end{array} \right\}, x > y \end{cases}, (x,y) \in [0,1] \times [0,1];$$

when $(q,p) \in [-1,1] \times (-\infty,0)$, we have

$$T_{(q,p)}(x,y) = \begin{cases} 1, x \leq y \;\; or \;\; q - x^p + y^p \leq 0 \\ (q - x^p + y^p)^{\frac{1}{p}}, otherwise \end{cases}, (x,y) \in [0,1] \times [0,1];$$

when $q \in [-1,1), p = 0$, we have

$$T_{(q,p)}(x,y) = \begin{cases} 1, x \leq y \\ y, x > y \end{cases}, (x,y) \in [0,1] \times [0,1];$$

when $q = 1, p = 0$, we have

$$T_{(q,p)}(x,y) = \begin{cases} 1, x \leq y \\ \frac{y}{x}, x > y \end{cases}, (x,y) \in [0,1] \times [0,1].$$

Proof. $T_{(q,p)}(x,y)$ is simply written as $\Rightarrow$. By $x \Rightarrow y = \sup\{z | x * z \leq y\}$ [1] we have the following.

1) Assume $x \leq y$. Then $x * z \leq x \leq y$. Hence $x \Rightarrow y = \sup\{z | 0 \leq z \leq 1\} = 1$.

2) Assume $x > y$. If $(q,p) \in [-1,0] \times (0,\infty)$, then

$$\begin{aligned} x \Rightarrow y &= \sup\{z | x * z \leq y\} \\ &= \sup\{z | x^p + z^p \leq |q|, 0 \leq y\} \vee \sup\{z | x^p + z^p \leq |q|, x \wedge z \leq y\} \\ &= [(|q| - x^p) \vee 0]^{\frac{1}{p}} \vee y; \end{aligned}$$

if $(q,p) \in [0,1] \times (0,\infty)$, then

$$\begin{aligned} x \Rightarrow y &= \sup\{z|x * z \le y\} \\ &= \sup\{z|x^p + z^p \le q, 0 \le y\} \vee \sup\{z|x^p > q, x \wedge z \le y\} \\ &\vee \sup\{z|x^p + z^p > q, x^p < q, (x^p + z^p - q)^{\frac{1}{p}} < y\} \\ &= \begin{cases} y, x^p > q \\ (q - x^p + y^p)^{\frac{1}{p}}, x^p \le q; \end{cases} \end{aligned}$$

if $(q,p) \in [-1,1] \times (0,\infty)$, then

$$\begin{aligned} x \Rightarrow y &= \sup\{z|x * z \le y\} \\ &= \sup\{z|xz = 0, 0 \le y\} \vee \sup\left\{z|xz \ne 0, (x^p + z^p - q)^{\frac{1}{p}} \le y\right\}. \\ &= \begin{cases} 1, q - x^p + y^p \le 0 \\ (q - x^p + y^p)^{\frac{1}{p}}, otherwise. \end{cases} \end{aligned}$$

For the case of $q \in [-1,1), p = 0$ *or* $q = 1, p = 0$, the conclusions of the theorem are evident.

The theorem is proved.

The implications in Theorem 3.2 are denoted by $\Rightarrow_{L\Pi GN(q,p)} (x,y), (x,y) \in [0,1]$ or $I_{L\Pi GN(q,p)}(x,y), (x,y) \in [0,1]$, and the following definition are given.

Definition 4. *The set* $\{I_{L\Pi GN(q,p)(q,p)}(x,y), (x,y) \in [0,1] | (q,p) \in [-1,1] \times (-\infty,+\infty)\}$ *is called the residual family of implication operators of* $T_{L\Pi GN(q,p)}$, *and it is denoted by* $I_{L\Pi GN(q,p)}, (q,p) \in [-1,1] \times (-\infty,+\infty)$, *or briefly* $I_{L\Pi GN(q,p)}$.

4 Triple I Method on FMP (FMT) Based on the Family $I_{(q,p)} - L\Pi GN$ of Implication Operators

Theorem 5. *The sustaining solution* $B^*(y)$ *of triple I based on* $I_{L\Pi GN(q,p)}$ *on FMP model is given by*

$$\begin{aligned} &B^*(y) = \sup_{x \in X} A^*(x) *_{L\Pi GN(q,p)} (I_{L\Pi GN(q,p)}(A(x), B(y))) \\ &= \sup_{\substack{(A^*(x))^p + I^p(A(x),B(y)) > |q| \\ xy \ne 0}} \{\lim_{t \to p} ((A^*(x))^t + I_{L\Pi GN(q,p)}{}^t(A(x),B(y)) - q)^{\frac{1}{t}} \\ &\wedge A^*(x) \wedge I_{L\Pi GN(q,p)}(A(x),B(y))\}, \\ &= \sup_{\substack{(A^*(x))^p + I^p{}_{L\Pi GN(q,p)}(A(x),B(y)) > |q| \\ xy \ne 0}} \{\lim_{t \to p} ((A^*(x))^t + I_{L\Pi GN(q,p)}{}^t(A(x),B(y)) - q)^{\frac{1}{t}} \\ &\quad \wedge A^*(x) \wedge I_{L\Pi GN(q,p)}(A(x),B(y))\}, \\ &y \in Y. \end{aligned}$$

Proof. From these papers [11-13] if implication I has residual left-continuous t-norm $*$, then the minimum fuzzy set $B^*(y) \in F(Y)$ in triple I sustaining method on FMP model by

$$B^*(y) = \sup_{x \in X} \{A^*(x) *_{L\Pi GN(q,p)} I_{L\Pi GN(q,p)}(A(x), B(y))\}, y \in Y,$$

we easily gained

$$B^*(y) = \sup_{\substack{(A^*(x))^p + I^p{}_{L\Pi GN(q,p)}(A(x),B(y)) > |q| \\ xy \neq 0}} \{\lim_{t\to p} ((A^*(x))^t + I_{L\Pi GN(q,p)}{}^t(A(x),B(y)) - q)^{\frac{1}{t}} \wedge A^*(x) \wedge I_{L\Pi GN(q,p)}(A(x),B(y))\}, y \in Y.$$

Theorem 6. *The sustaining solution $A^*(y)$ of triple I based on $I_{L\Pi GN(q,p)}$ on FMT model is given by*

$$A^*(x) = \inf_{x\in X} \left\{ I_{L\Pi GN(q,p)}(I_{L\Pi GN(q,p)(A(x),B(y))},\ B^*(y)) \right\},$$

$$= \inf_{I_{L\Pi GN(q,p)}(A(x),B(y)) > B^*(y)} \left\{ I_{L\Pi GN(q,p)}(I_{L\Pi GN(q,p)(A(x),B(y))},\ B^*(y)) \right\}, x \in X$$

Proof. It is early to prove the theorem from the paper [13].

5 Conclusion and Expectation

The paper mainly gives the family of t-norms $*_{L\Pi GN(q,p)}$ satisfying

$$x *_{L\Pi GN(q,p)} y = \begin{cases} \lim\limits_{t\to p} (x^t + y^t - q)^{\frac{1}{t}} \cap x \cap y, x^p + y^p > |q|, xy \neq 0 \\ 0, \text{otherwise} \end{cases}, (x,y) \in [0,1]\times[0,1]$$

(where $q \in [-1,1], p \in (-\infty,+\infty)$), the residual family $\Rightarrow_{L\Pi GN(q,p)}$ of implication operators of $*_{L\Pi GN(q,p)}$ as follows satisfies that

$$x \Rightarrow_{L\Pi GN(q,p)} y = \begin{cases} 1, x \leq y \\ (1 - f + f.y) \vee \lim\limits_{t\to p} ((|q| - x^t) \vee 0)^{\frac{1}{|t|}} f, q - x^p + y^p \leq 0, x > y \\ \lim\limits_{t\to p^+} (q - x^t + y^t)^{\frac{1}{t}} \vee f.y \vee \lim\limits_{t\to p} ((|q| - x^t) \vee 0)^{\frac{1}{|t|}} f, q - x^p + y^p > 0, x > y \end{cases}$$

(where $t = 0, 0^t = 0^{\frac{1}{t}} = 0$), and solution of triple I method based on this family of implication operators as follows:

$$B^*(y) = \sup_{\substack{(A^*(x))^p + I^p{}_{L\Pi GN(q,p)}(A(x),B(y)) > |q| \\ xy \neq 0}} \{\lim_{t\to p} ((A^*(x))^t + I_{L\Pi GN(q,p)}{}^t(A(x),B(y)) - q)^{\frac{1}{t}} \wedge A^*(x) \wedge I_{L\Pi GN(q,p)}(A(x),B(y))\}, y \in Y$$

and

$$A^*(x) = \inf_{I_{L\Pi GN(q,p)}(A(x),B(y)) > B^*(y)} \left\{ I_{L\Pi GN(q,p)}(I_{L\Pi GN(q,p)(A(x),B(y))},\ B^*(y)) \right\}, x \in X.$$

The above conclusions have important significance since on the one hand it affords many new implication operators; on the other hand it is convenient to optimize implication operator in fuzzy reasoning by the a uniform form.

Acknowledgments. This work was supported by National Natural Science Foundation of China Grant No. 60874067.

References

1. Hájek, P.: Metamathematics of Fuzzy Logic. Kluwer Academic, Lodon (1998)
2. Zádeh, L.: Fuzzy sets. Information and Control 8, 338–353 (1965)
3. Zádeh, L.A.: Outline of a new approach to the analysis of complex systems and decision processes. IEEE Trans. Syst. Man Cyber. 3, 28–44 (1973)
4. Fodor, J.C., Roubens, M.: Fuzzy Preference Modelling and Multicriteria Decision Support. Kluwer Academic Publishers, Dordrecht (1994)
5. Schweizer, B., Sklar, A.: Statistical metric spaces. Pacific J. Math. 10, 313–334 (1960)
6. Schweizer, B., Sklar, A.: Probabilistic Metric Spaces. North-Holland, New York (1983)
7. Menger, K.: Statistical metrics. Proc. Nat. Acad. Sci., USA 8, 535–537 (1942)
8. Klement, E.P., Weber, S.: Generalized measures. Fuzzy Sets and Systems 40, 375–394 (1991)
9. Butnariu, D., Klement, E.P.: Triangular Norm-Based Measures and Games with Fuzzy Coalitions. Kluwer Academic Publishers, Dordrecht (1993)
10. Frank, M.J.: On the simultaneous associativity of $F(x, y)$ and $x + y - F(x, y)$. Aequationes Math. 19, 194–226 (1979)
11. Wang, G.J.: Fully implicational triple I method for fuzzy reasoning. Sci. China (Ser. E) 29, 43–53 (1999) (in Chinese)
12. Wang, G.J.: Introduction to mathematical logic and resolution principle, 2nd edn. Science Press, Beijing (2009)
13. Pei, D.W.: Unified full implication algorithms of fuzzy reasoning. Information Sciences 178, 520–530 (2008)
14. Jeneia, S., Montagnab, F.: Ageneral method for constructing left-continuous t-norms. Fuzzy Sets and Systems 136, 263–282 (2003)
15. Whalen, T.: Parameterized R-implications. Fuzzy sets and systems 134, 231–281 (2003)
16. Zhang, X.H., He, H.C., Xu, Y.: Fuzzy logic system of based on Schweizer-Sklar T-norm. Sci. China (Ser. E) 35, 1314–1326 (2006) (in Chinese)
17. Zhang, X.F., Meng, G.W., Zhang, A.Y.: The Families of Implication Operators and Their Application. Chinese Journal of Computers 30, 448–453 (2007) (in Chinese)

The Properties of Normal R_0 Algebras

Hua-Rong Zhang

School of Science, China Jiliang University,
Hangzhou 310018, China
e-mail: hrzhang2008@cjlu.edu.cn

Abstract. In this paper, we consider the properties of normal R0 algebras and the relationship between normal R0 algebras and other algebras. We also discuss the filters of normal R0 algebras. We get that in a R0 algebra, the following conditions are equivalent : F is an implicative filter of L; F is a positive implicative filter of L; F is a Boolean filter of L. And F is a filter if and only if F is a fantastic filter.

Keywords: Normal R0 algebra, BL algebra, Boolean algebra, MV algebra, Filter.

1 Introduction

With the development of mathematics and computer science, non-classical logic has been extensively studied. So far, many-valued logic has become an important part of non-classical logic. In order to research the logical system whose propositional value is given in a lattice from the sematic viewpoint, Xu [1] proposed the concept of lattice implication algebras and discussed some of their properties. Xu and Qin [2] introduced the notion of implicative filters in a lattice implication algebra. BL algebras have been invented by P. Hajek [3] in order to provide an algebraic proof of the completeness theorem of “Basic Logic” (BL ,for short) . It is arisen from the continuous triangular norms, familiar in the fuzzy logic framework. Filters in BL algebras are also defined. R0 algebras have been introduced by Wang [4] in order to provide an algebraic proof of the completeness theorem of a formal deductive system. Pei [5] studied the filters of R0 algebras. Note that R0 algebras are different from BL algebras because the identity $x \wedge y = x \otimes (x \to y)$ holds in BL algebras, and does not hold in R0 algebras. R0 algebras are also different from lattice implication algebras because the identity $(x \to y) \to y = (y \to x) \to x$ holds in lattice implication algebras and does not hold in R0 algebras. In these algebras, all kinds of filters such as implicative filters, positive implicative filters, fantastic filters, Boolean filters are introduced. Corresponding properties are discussed. In recent years, a great deal of literature has been produced on the theory of filters and fuzzy filters. In [6], fantastic filters are introduced into R0 algebras. In [7], it is proved that if F is a normal MP-filter of R0 algebra L, then $L/\sim_F$ is a normal R0 algebra. Based on this, in this paper, we discuss the properties of normal R0 algebras and study the relationship between normal R0 algebras and other algebras.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 177–185.
springerlink.com

2 Preliminaries

Definition 2.1. [4] *Let L be a bounded distributive lattice with order-reversing involution* $\neg$ *and a binary operation* $\rightarrow$. *(L;* $\neg;\vee;\rightarrow$ *) is called a R0 algebra if it satisfies the following axioms:*

(1) $\neg x \rightarrow \neg y = y \rightarrow x$,
(2) $1 \rightarrow x = x, x \rightarrow x = 1$,
(3) $y \rightarrow z \leq (x \rightarrow y) \rightarrow (x \rightarrow z)$,
(4) $x \rightarrow (y \rightarrow z) = y \rightarrow (x \rightarrow z)$,
(5) $x \rightarrow (y \vee z) = (x \rightarrow y) \vee (x \rightarrow z)$,
(6) $(x \rightarrow y) \vee ((x \rightarrow y) \rightarrow \neg x \vee y) = 1$.

Let L be a R0 algebra. Define $x \otimes y = \neg(x \rightarrow \neg y)$, *for any* $x, y \in L$. *It is proved that* $(L, \wedge, \vee, \otimes, \rightarrow, 0, 1)$ *is a residual lattice.*

Definition 2.2. [8] *By an NM-algebra is meant a structure* $(L, \vee, \wedge, \otimes, \rightarrow, 0, 1)$ *of type (2, 2, 2, 2, 0, 0) such that for all* $x, y, z \in L$,

(1) $(L, \wedge, \vee, 0, 1)$ *is a bounded lattice,*
(2) $(L, \otimes, 1)$ *is a monoid,*
(3) $x \otimes y \leq z$ *iff* $x \leq y \rightarrow z$,
(4) $(x \rightarrow y) \vee (y \rightarrow x) = 1$,
(5) $((x \otimes y) \rightarrow 0) \vee ((x \wedge y) \rightarrow (x \otimes y)) = 1$,
(6) $(x \rightarrow 0) \rightarrow 0 = x$.

In [9], Pei proved that R0-algebras and NM-algebras are the same algebraic structures.

Example 2.1. [4] Let L= [0, 1]. For any $x, y \in$ L, we define:

$$x \wedge y = \min\{x, y\} , \ x \vee y = \max\{x, y\} ,$$
$$\neg x = 1 - x, x \rightarrow y = \begin{cases} 1 & x \leq y \\ \neg x \vee y & x > y \end{cases} ,$$

Then (L; $\neg$; $\vee$; $\rightarrow$) is a R0 algebra.
In what follows, L will denote a R0 algebra, unless otherwise specified.

Lemma 2.1. [4] For any $x, y \in$ L, the following properties hold:

(1) $0 \rightarrow x = 1, x \rightarrow 0 = \neg x, x \rightarrow x = 1, x \rightarrow 1 = 1,$
(2) $x \leq y$ implies $y \rightarrow z \leq x \rightarrow z, z \rightarrow x \leq z \rightarrow y,$

(3) $x \to y \le (y \to z) \to (x \to z)$,
(4) $((x \to y) \to y) \to y = x \to y$,
(5) $x \to (y \wedge z) = (x \to y) \wedge (x \to z)$,
(6) $(x \vee y) \to z = (x \to z) \wedge (y \to z)$,
(7) $(x \wedge y) \to z = (x \to z) \vee (y \to z)$,
(8) $x \le y$ if and only if $x \to y = 1$,
(9) $x \vee y = ((x \to y) \to y) \wedge ((y \to x) \to x)$.

Definition 2.3. [3] *A BL algebra is an algebra* $(L, \wedge, \vee, \otimes, \to, 0, 1)$ *with four binary operations* $\wedge, \vee, \otimes, \to$, *and two constants 0,1 such that:*

(1) $(L; \wedge; \vee; 0; 1)$ *is a bounded lattice,*
(2) $(L; \otimes; 1)$ *is a commutative monoid,*
(3) $\otimes$ *and* $\to$ *form an adjoint pair, i.e.,*

$$x \otimes y \le z \text{ iff } x \le y \to z \text{, for all } x, y, z \in L,$$

(4) $x \wedge y = x \otimes (x \to y)$,
(5) $(x \to y) \vee (y \to x) = 1$.

A BL algebra is called an MV algebra if

$$(y \to x) \to x = (x \to y) \to y \text{, for all } x, y \in L.$$

Definition 2.4.[7] *A R0-algebra L is called a normal R0-algebra if it satisfies:*

$$(y \to x) \to x = (x \to y) \to y \text{, for all } x, y \in \mathrm{L}.$$

3 The Existence of Normal R0 Algebras and Their Properties

Definition 3.1 [5]. *A non-empty subset F of a R0-algebra L is called a filter of L if it satisfies:*

(F1) $1 \in F$;
(F2) $x \in F$ *and* $x \to y \in F$ *imply* $y \in F$ *for all* $x, y \in L$.

Definition 3.2. *A non-empty subset F of a R0-algebra L is said to be a fantastic filter of L if it satisfies:(F1) and (F3) :*

$z \to (y \to x) \in F$ *and* $z \in F$ *imply* $((x \to y) \to y) \to x \in F$ *for all* $x, y, z \in L$.

The following example shows that the fantastic filters of R0 algebras exist.

Example 3.1. [7] *Let L be the chain {0, a,b,c ,1} with Cayley tables as follows:*

x	$\neg x$
0	1
a	c
b	b
c	a
1	0

$\to$	0	a	b	c	1
0	1	1	1	1	1
a	c	1	1	1	1
b	b	b	1	1	1
c	a	a	b	1	1
1	0	a	b	c	1

Define $\vee$ operations and $\wedge$ operations on L by $x \vee y = \max\{x, y\}$ and $x \wedge y = \min\{x, y\}$ for all $x, y \in$ L. By routine calculation , we can obtain that L is a R0-algebra and F = {c, 1} is a normal MP-filter of L .

Theorem 3.1. [7] *Let F be a filter of a R0 algebra L. Then F is a normal MP-filter if and only if $L/\sim_F$ is a normal R0 algebra.*

In [6], normal MP-filters are called fantastic filters.

By Theorem 3.1 , we get in example3.1, $L/\sim_F = \{[1]_F, [b]_F, [0]_F\}$ is a normal R0 algebra. This shows the existence of normal R0 algebras.

Theorem 3.2. *In each normal R0 algebra, the following relations hold, for all $x, y, z \in L$,*

(1) $x \vee y = (y \to x) \to x = (x \to y) \to y$

(2) $x \wedge y = x \otimes (x \to y)$

(3) $((x \to y) \to y) \to x = y \to x$

Proof. (1) Since

$$x \vee y = ((y \to x) \to x) \wedge ((x \to y) \to y)$$

and

$$(y \to x) \to x = (x \to y) \to y ,$$

we have

$$x \vee y = (y \to x) \to x = (x \to y) \to y .$$

(2) Since

$$x \otimes (x \to y) = \neg(x \to \neg(x \to y)) ,$$

we need to show

$$x \to \neg(x \to y) = \neg(x \wedge y) ,$$

that is

$$x \to \neg(x \to y) = \neg x \vee \neg y \ .$$

By (1), we have

$$\neg x \vee \neg y = (\neg y \to \neg x) \to \neg x = x \to \neg(x \to y) \ .$$

The proof is complete.
(3) By Lemma 2.1, we have

$$(y \to x) \to (((x \to y) \to y) \to x \ = ((x \to y) \to y) \to ((y \to x) \to x) = 1,$$

hence

$$y \to x \le ((x \to y) \to y) \to x \ .$$

Conversely,

$$(((x \to y) \to y) \to x) \to (y \to x) \ge y \to ((x \to y) \to y) \ = (x \to y) \to (y \to y) = (x \to y) \to 1 = 1$$

Therefore

$$y \to x \ge ((x \to y) \to y) \to x \ .$$

The proof is complete.
By the properties of R0 algebras and Theorem 3.2, we have

Theorem 3.3. *Every normal R0 algebra is a MV algebra.*

Corollary 3.1. *Every normal R0 algebra is a BL algebra.*

Theorem 3.4. *Every Boolean algebra is a normal R0 algebra.*

Proof. Firstly, we have that every Boolean algebra L is a R0 algebra[4].Secondly, in [4], define $x \to y = x' \vee y$, therefore

$$(x \to y) \to y = x' \vee y \to y = (x' \vee y)' \vee y$$
$$= (x \wedge y') \vee y = (x \vee y) \wedge (y' \vee y) = (x \vee y) \wedge 1 = x \vee y \ .$$

Similarly, we have

$$(y \to x) \to x = x \vee y \ .$$

Hence

$$(x \to y) \to y = y \to (y \to x) \ .$$

It is showed that every Boolean algebra is a normal R0 algebra.

Corollary 3.2. *Every Boolean algebra is a MV algebra.*

Lemma 3.1. *In a normal R0 algebra, the following conditions are equivalent: for all* $x, y \in \mathrm{L}$,

(1) $(x \to y) \to x = x$

(2) $x \wedge \neg x = 0$ where $\neg x = x \to 0$

(3) $x \vee \neg x = 1$

(4) $x = \neg x \to x$

Proof. ($1 \Rightarrow 3$) By (1),

$$(x \to 0) \to x = x\ ,$$

we have

$$x \vee \neg x = (\neg x \to x) \to x = x \to x = 1\ .$$

($3 \Rightarrow 4$) Since

$$x \vee \neg x = 1,$$

we have

$$(\neg x \to x) \to x = 1,$$

hence

$$\neg x \to x \leq x.$$

And

$$x \leq \neg x \to x,$$

therefore

$$x = \neg x \to x\ .$$

($4 \Rightarrow 1$) By Lemma 2.1,

$$x \to ((x \to y) \to x) = (x \to y) \to (x \to x) = 1\ ,$$

hence

$$x \leq (x \to y) \to x.$$

By Lemma 2.1, we have

$$x \to y \geq x \to 0\ ,$$

$$(x \to y) \to x \leq (x \to 0) \to x\ ,\ ((x \to y) \to x) \to x \geq ((x \to 0) \to x) \to x = x \to x = 1\ ,$$

therefore

$$(x \to y) \to x \leq x\ .$$

Hence

$$(x \to y) \to x = x\ .$$

($2 \Leftrightarrow 3$) $x \wedge \neg x = 0$ if and only if $\neg(x \wedge \neg x) = 1$ if and only if $\neg x \vee x = 1$.

Theorem 3.5. *Let L is a normal R0 algebra, L is a Boolean algebra if and only if* $(x \to y) \to x = x$ *, for all* $x, y \in L$*.*

Proof. It is clear.

4 The Filters of Normal R0 Algebras

In a normal R0 algebra, we can introduce the notions of filters, implicative filters, positive implicative filters, Boolean filters, fantastic filters.

Definition 4.1. *Let L be a normal R0 algebra. A subset F of L is called a filter of L if it satisfies: (F1) and (F2).*

Definition 4.2. *Let L be a normal R0 algebra. A subset F of L is called an implicative filter if it satisfies:*

(F1) and (F4): $x \to (y \to z) \in F$ *and* $x \to y \in F$ *imply* $x \to z \in F$*.*

Definition 4.3. *Let L be a normal R0 algebra. A subset F of L is called a Boolean filter if* $x \vee \neg x \in F$ *for all* $x \in L$*.*

Definition 4.4. *Let L be a normal R0 algebra. A subset F of L is called a positive implicative filter if it satisfies:*
(F1) and (F5): $x \to ((y \to z) \to y) \in F$ *and* $x \in F$ *imply* $y \in F$*, for all* $x, y, z \in L$*.*

Definition 4.5. *A non-empty subset F of an R0-algebra L is said to be a fantastic filter of L if it satisfies: (F1) and (F3) .*

Theorem 4.1. In a R0 algebra, the following conditions are equivalent:

(1) F is an implicative filter of L;
(2) F is a positive implicative filter of L;
(3) F is a Boolean filter of L.

Corollary 4.1. In a normal R0 algebra, the following conditions are equivalent:

(1) F is an implicative filter of L;
(2) F is a positive implicative filter of L;
(3) F is a Boolean filter of L.

Theorem 4.2. *Let L be a normal R0 algebra. F is a filter of L if and only if F is a fantastic filter of L.*

Proof. Let

$$z,\ z \to x \in \mathrm{F}.$$

Since

$$z \to (1 \to x) = z \to x \in \mathrm{F}$$

and F is a fantastic filter, then

$$((x \to 1) \to 1) \to x = x \in \mathrm{F},$$

this shows that F is a filter.

Conversely, let F is a filter and

$$z \to (x \to y) \in \mathrm{F},\ \ z \in \mathrm{F}.$$

Then

$$x \to y \in \mathrm{F}.$$

By Theorem3.2, we have

$$((y \to x) \to x) \to y \in \mathrm{F}.$$

The proof is complete.

5 Conclusion

We discuss the properties of normal R0 algebras ,study the relationship between normal R0 algebras and other algebras such as MV algebras, Boolean algebras, BL algebras .We introduce the notion of filters, implicative filters, positive implicative filters, fantastic filters into normal R0 algebras. We get that in a normal R0 algebra, the following conditions are equivalent: F is an implicative filter of L; F is a positive implicative filter of L; F is a Boolean filter of L; and F is a filter if and only if F is a fantastic filter.

References

1. Xu, Y.: Lattice implication algebras. J. Southwest Jiaotong Univ. 1, 20–27 (1993)
2. Xu, Y., Qin, K.Y.: On filters of lattice implication algebras. J. Fuzzy Math. 2, 251–260 (1993)
3. Hajek, P.: Metamathematics of fuzzy logic. Kluwer Academic Publishers, Dordrecht (1998)
4. Wang, G.J.: Non-classical mathematical logic and approximate reasoning. Science Press, Beijing (2000)

5. Pei, D.W.: Implication filters and congruence relations in R0 algebras. Xi'an United University 13(4), 24–29 (2000)
6. Zhang, H.R.: The filters theory of R0 algebras. Fuzzy System and Mathematics 3, 17–21 (2009)
7. Liu, Y.L., Ren, M.Y.: Normal MP-Filters of R0-Algebras. Fuzzy Info. and Engineering, ASC 54, 113–118 (2009)
8. Esteva, F., Godo, J.: Monoidal t-norm-based logic: towards a logic for left-continuous t-norms. Fuzzy Sets and Systems 124, 271–288 (2001)
9. Pei, D.W.: On equivalent forms of fuzzy logic systems NM and IMTL. Fuzzy Sets and Systems 138, 187–195 (2003)

A Splitting Algorithm Based on Soft Constraints for the Computation of Truth Degree in Quantitative Logic

Bang-he Han[1,2] and Yong-ming Li[1,3]

[1] College of Mathematics and Information Science, Shaanxi Normal University, Xi'an, 710062, China
e-mail: bhhan@stu.snnu.edu.cn

[2] School of Science, Xidian University, Xi'an, 710071, China

[3] College of Computer Science, Shaanxi Normal University, Xi'an, 710062, China
e-mail: liyongm@snnu.edu.cn

Abstract. The concept of the truth degree of a formula is the crucial tool and the building block in quantitative logic. So how to compute the truth degree of a formula efficiently is a principal question in this subject. This paper aims at bringing an optimal method for doing this job. Firstly, characterizations of the concept of truth degree are made with concepts from soft constraint theory. Particularly, two soft constraint systems are proposed such that formulas can be taken as soft constraints over them. Then by exploiting the algebraic properties of both constraint systems and n-valued propositional systems, it is shown that different soft constraint systems plays different roles in the computation of truth degree of formulas. An optimal method named splitting algorithm for computing truth degrees of formulas is proposed.

Keywords: Truth Degree, Quantitative Logic, Soft Constraint, Model Counting, Constraint Solving, Splitting Algorithm.

1 Introduction

In the real world, not every question can be answered by just yes or no. For instance, when a student has one leg in the classroom with the other out, then is the student in the classroom? Zadeh [1] proposed the idea of fuzzy sets by which the classical concept of set is fuzzified. That is: every element of the domain is associated a degree belong to [0,1], which characterizes the extent to which this element can be regarded as an element of the set. We call this method *a graded approach*. Now we can answer the question above: if most of the body is in the classroom, then maybe we can say that this student is in the classroom with 90 percent.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 187–198.
springerlink.com

The using of *graded approach* is an important and useful tool in uncertainty-based reasoning and also some other subjects. For instances, (1) Pavelka [2] achieved wonderful results on fuzzy logic by grading the inferencing rules and the proof process with lattice value. (2) The solution of classical satisfaction problems(CSPs) is to find the values of variables such that all constraints are satisfied. However, in daily life, people may have different preferences for different choice. By using *a graded approach*, soft constraint satisfaction problems (SCSPs)have been proposed to characterize constraint models with criteria such as preferences, costs or priorities by Bistarelli [3]. SCSP is more representable and complicated that CSP.

In this paper, we want to investigate another research work which provide a graded approach to propositional logic. It is well known that mathematical logic is the formalized theory with the character of symbolization and it lays stresses on formal deduction rather than on numerical computation. On the contrary, numerical computation aims to solve various computing problems by means of possible methods such as interpolation, iteration, difference, probability estimation, etc., and numerical computation pays close attention to problem-solving as well as to error estimation but seldom uses formal deduction. Hence, mathematical logic and numerical computation are two branches of mathematics miles apart. A new branch quantitative logic, which is the result of combining together mathematical logic and probability computation, has been proposed by Wang [4-5].

Wang [4-5] proposed the concept of the degree of the truth in the framework of so called many-valued propositional logic systems with the intention of measuring to what extent a given formula is true. Such a concept can indeed induce, in a very natural way, the degree of the similarity as well as to induce a pseudo-metric among formulas as the graded version of the notion of the logical equivalence. The basic properties of such induced logic metric space hence are investigated. What's more, different concepts of the degree of the divergence and the degree of the consistency in order to grade the extent of the consistency of a logic theory were given. To the end, all the basic logic notions are graded. Lastly, three patterns of approximate reasoning so far as the many-valued propositional logic systems are proposed . Since this theory has touched upon several key concepts of similarity, comparison, measurements, etc. in information theory, the theory may have profound impact in contemporary information technology. In addition, quantitative logic can be regarded as a highly representative example of mathematics of uncertainty capable of handling vagueness.

In recent years many researchers have made contributions to quantitative logic: Zhou [6] discussed the consistency degree of theories in quantitative logic; Han [7] proposed the concept of conditional truth degree in classical logic based on the idea of conditional probability; Wang, et.al[8] brought randomized theory truth degree into quantitative logic; Zhang [9] pointed out a syntactic graded method of two-valued propositional logic formulas. A concept of absolute truth of formulas in n-valued Lukasiewicz propositional logic was introduced by [10]; Wang, et al [11] developed the theory of quantitative logic into modal logic; Han [12-13] discussed the problems of error accumulation and convergency theory in quantitative logic.

The concept of the truth degree of a formula is the crucial tool and the building block in quantitative logic. All the concepts in quantitative logic is based on the concept of truth degree of formulas. Theoretically, the truth degrees of formulas are well-defined. Particularly, it is easy to show that the computation of the truth degree of a formula in classical logic is equivalent to counting the models of this formula. But in general the given definition does not yield a practical algorithm to compute these degrees. As far as the authors know, none of the progresses in quantitative logic so far concerns how to compute the truth degree of a formula in efficient ways. We think that this is an obstacle to the development of quantitative logic. Nevertheless, there do exist optimal methods in other research fields that can be borrowed for computing the truth degree of formulas. A number of different techniques for model counting have been proposed over the last few years. For example, Relsat by Bayardo [14] extends systematic SAT solvers for model counting and uses component analysis for efficiency, Cachet by Sang et al, [15] adds caching schemes to this approach, c2d by Darwiche [16] converts formulas to the dDNNF form which yields the model count as a byproduct, ApproxCount by Wei and Selman [17] and Sample Count by Gomes, et al [18] exploit sampling techniques for estimating the count, MBound by Gomes, et al, [19] relies on the properties of random parity or xor constraints to produce estimates with correctness guarantees, and the recently introduced Sample Minisat by Gogate and Dechter [20] uses sampling of the backtrack free search space of systematic SAT solvers. BPCount and MiniCount by Kroc [21] provdes useful information even on structure loopy formulas. Samer and Szeider [22] brought in algorithms based on tree-decomposition of graphs associated with the given CNF formula, in particular primal, dual, and incidence graphs. Favier, et al [23] propose to adapt BTD for solving the CSP problem(thus also for SAT problem.

However, this paper's contribution is not just an easy shifting, i.e., we don't make just a list of the algorithms already exists. Since we think that the methods already existed lay emphasis on how to compute but lack of concerning about the underlying essence: why how. In order to make up this deficiency, we start from the underlying algebraic property of the corresponding systems. Then based on this property, we propose a splitting algorithm by which the truth degree of a propositional formula can be presented in a mathematical expression rather than a program.

The remainder of this paper is organized as follows. Section 2 introduces the basic theory needed in this paper including quantitative logic, soft constraint system. In Section 3, we make characterizations of the concept of truth degree with concepts from soft constraint systems and an algorithm named splitting algorithm is proposed. Section 4 is the conclusion and discussion part.

2 Preliminaries

In this section, we first present a brief introduction on the concepts of truth degree of a formula in n-valued propositional logic systems. Then the theory of soft constraint satisfaction problems (SCSPs) is also briefly reviewed.

2.1 The Concept of Ttruth Degree of Formulas (Wang [4-5])

In this subsection, we assume that readers are familiar with the common n-valued propositional logic systems: classical logic L, Łukasiewicz n-valued system Łn, $\mathscr{L}_n^*$ [5], Gödel n-valued system G_n and Product n-valued system $\prod_n$. S is the set of axiomatic formulas, and $F(S)$ is the set of all formulas. $W_n = \{0, 1/(n-1), \cdots, (n-2)/(n-1), 1\}$ is the corresponding valuation domain, recall that W_n is an algebra of type $(\neg, \vee, \rightarrow)$. For example, in system Łn, $\rightarrow$ means the Łukarsiwicz implication operator $\rightarrow_L$. In n-valued propositional logic, we write $A \otimes B$ to represent the formula $\neg(A \rightarrow \neg B)$, the semantics of $\otimes$ corresponds to the t-norm $\otimes_L$, $\otimes_o$, $\otimes_G$, $\otimes_\pi$. Let $A = A(p_1, \cdots, p_m)$ be a formula. By substituting $x_1, \cdots, x_m$ for $p_1, \cdots, p_m$, respectively, and interpreting the logical connectives $\neg, \vee, \rightarrow$ as the corresponding operations $\neg, \vee, \rightarrow$ on W_n, we then obtain an m-ary function $\bar{A}(x_1, \cdots, x_m) : W_n^m \rightarrow W$, called the truth function induced by A.

Definition 1. *(Truth degree) Let $A = A(p_1, \cdots, p_m)$ be a formula containing m atomic formulas $p_1, \cdots, p_m$ in a certain n-valued propositional system, and let $\bar{A}$ be the truth function induced by A. Define*

$$\tau_n(A) = \frac{1}{n^m} \sum_{i=1}^{n-1} \frac{i}{n-1} \mid \bar{A}^{-1}(\frac{i}{n-1}) \mid,$$

where $|E|$ denotes the number of elements of the set E. $\tau_n(A)$ is called the degree of the truth of A in the n-valued system. Particularly, in the case of $n = 2$,

$$\tau_2(A) = \frac{\mid \bar{A}^{-1}(1) \mid}{2^m}.$$

There are several kinds of the concept of truth degree. For example, a concept of absolute truth of the latter in n-valued Łukasiewicz propositional logic was introduced in [10].

Definition 2. *(Absolute Truth Degree) Let $A = A(p_1, \cdots, p_m)$ be a formula containing m atomic formulas $p_1, \cdots, p_m$ in a certain n-valued propositional system, and let $\bar{A}$ be the truth function induced by A. Define the absolute truth degree of A, denoted by $\theta_n(A)$, as follows:*

$$\theta_n(A) = \frac{1}{n^m} \mid \bar{A}^{-1}(1) \mid .$$

2.2 Soft Constraint Satisfaction Problems (SCSPs)

The framework of soft constraint satisfaction problems is based on a semiring structure, where the set of the semiring specifies the values to be associated with each tuple of values of the variable domain, and the two semiring operations $+, \times$ model constraint projection and combination respectively. We call a tuple $S = (A, +, \times, 0, 1)$ a semiring, if A is a set containing different elements 0 and 1, and $+, \times$ are operations on S satisfying the following properties: $+$ is associative and commutative with identity 0, $\times$ is associative with identity 1 and null element 0, (i.e., for all $a \in S, a \times 0 = 0 \times a = 0$), and $\times$ distributes over $+$, i.e., for all $a, b, c \in S$, $a \times (b + c) = (a + b) \times (a + c)$ and $(b + c) \times a = (b + a) \times (c + a)$. A c-semiring is a semiring such that $+$ is idempotent (i.e., for all $a \in S$, $a + a = a$), $\times$ is commutative, and 1 is the absorbing element of $+$.

Example 1. (i) Let R^+ be the set of nonnegative real numbers, $+, \times$ are the natural addition and multiplication operations. Then $< R^+, +, \times, 0, 1 >$ is a semiring. This semiring is often used to count numbers or computing the sum of a certain set of real numbers. Note that $< R^+, +, \times, 0, 1 >$ is not a c-semiring.

(ii) The algebraic structure $([0, 1], \vee, \otimes, 0, 1)$ is a c-semiring, where $\otimes$ means a t-norm on $[0, 1]$.

Definition 3. *(Soft Constraint Systems) Suppose $S =< A, +, \times, 0, 1 >$ is a semiring. V is a set of variables. The frames of variables are same, denoted by D. We call the tuple $CS =< S, D, V >$ a soft constraint system. A constraint c is a pair $< def, con >$, where*

(1) $con \subseteq V$, it is called the type of the constraint;

(2) $def : D^k \to A$ (where k is the cardinality of con*)is called the value of the constraint.*

Definition 4. *(Combination Operation) Given a soft constraint system $CS =< S, D, V >$, where $S =< A, +, \times, 0, 1 >$, and two constraints $c_1 =< def_1, con_1 >$ and $c_2 =< def_2, con_2 >$ over CS, their combination, written as $c_1 \otimes c_2$, is the constraint $c =< def, con >$ with*

$$con = con_1 \cup con_2,$$

$$def(t) = def_1(t \downarrow^{con}_{con_1}) \times def_2(t \downarrow^{con}_{con_2}), t \in D^{|con|}.$$

Definition 5. *(Projection Operation) Given a soft constraint system $CS =< S, D, V >$, where $S =< A, +, \times, 0, 1 >$, and a constraint $c =< def, con >$ over CS, and a set I of variables, the projection of c over I, written as $c \Downarrow_I$, is the constraint $c =< def', con' >$ over CS with*

$$con' = I \cap con,$$

$$def'(t') = \sum_{\{t | t \downarrow^{con}_{I \cap con} = t'\}} def(t), t' \in D^{|con'|}.$$

Definition 6. *(SCSP Problem and its solution) Given a constraint system* $CS =< S, D, V >$*, where* $S =< A, +, \times, 0, 1 >$*, a constraint problem over* CS *is a pair* $P =< C, con >$*, where* C *is a set of constraints over* CS *and* $con \subseteq V$*. The solution of* P *is defined as the constraint* $(\otimes C) \Downarrow_{con}$*, denoted by* $Sol(P)$*.*

Definition 7. *(Best level of consistency) Given a constraint system* $CS =< S, D, V >$*, where* $S =< A, +, \times, 0, 1 >$*, a constraint problem over* CS *is a pair* $P =< C, con >$ *over* $CS =< S, D, V >$ *define* $blevel(P) \in S$ *by* $< blevel(P), \emptyset >= (\otimes C) \Downarrow \emptyset$*.*

Theorem 1. *Given a constraint system* $CS =< S, D, V >$*, where* $S =< A, +, \times, 0, 1 >$*, and a constraint problem* $P =< C, con >$*, where* C *contains two constraints* $c_1 =< def_1, con_1 >$ *and* $c_2 =< def_2, con_2 >$*,if* $con_1 \cap con_2 \subseteq con \subseteq con_1 \cup con_2$*, then we have*

$$(c_1 \otimes c_2) \Downarrow_{con} = c_1 \Downarrow_{con \cap con_1} \otimes c_2 \Downarrow_{con \cap con_2} .$$

For the lack of space, we omit the proof here.

3 A Splitting Algorithm for Computing the Truth Degree of Formulas in Quantitative Logic

3.1 Characterizations of "Truth Degree" with Concepts from Constraint Theory

Let S be the semiring $< R^+, +, \times, 0, 1 >$, V is the set of atomic propositions, and $D = W_n$. Then we get a soft constraint system $< R^+, D, V >$ with respect to the underlying n-valued propositional logic systems. Similarly, when take the semiring as $([0, 1], \vee, \otimes, 0, 1)$, where $\otimes$ corresponds to the underlying t-norm in the semantics of the n-valued propositional logic systems, D, V remain the same, we can get another soft constraint system$< [0, 1], D, V >$. We will show that different soft constraint system has different effect on the computation problem in Ql.

Suppose $A = A(p_1, \cdots, p_m)$ be a formula in $F(S)$ containing m atomic formulas $p_1, \cdots, p_m$, i.e., $var(A) = \{p_1, ..., p_m\}$ and let $\bar{A}$ be the truth functions induced by A, then:

Proposition 1. *(i)* A *can be taken as a constraint of the soft constraint system*$< R^+, +, \times, 0, 1 >$*(or* $< [0, 1], D, V >$ *), denoted by* $c_A =< def_A, con_A >$*, where* $def_A = \bar{A}$*,* $con_A = var(A)$*.*

(ii) Suppose Γ *be a theory (i.e., a set of formulas)in* $F(S)$*, then we can take* Γ *as a set of constraints. In this paper, we call* c_A *a propositional constraint induced by* A*.*

Proposition 2. *Let* c_A, c_B *be two propositional constraints induced by* A, B *over the constraint system* $< [0, 1], D, V >$*. Then the definition function* $def_{A\otimes B}$ *of* $c_{A\otimes B}$ *is equal to* $def_A \otimes def_B$*.*

Proposition 3. *Let c_A be a propositional constraint induced by A over the constraint system $< R^+, D, V >$. Consider the soft constraint problem $P =< \{c_A\}, con >$, where $con \subseteq var(A)$, then $\forall\ t \in \Omega_{con}$, $c_A \Downarrow_{con}(t)$ counts the cardinality of the set $\{\bar{A}(t^{'})|t^{'} \in \Omega_{con_A}, t^{'} \downarrow^{con_A}_{con}= t\}$.*

Proposition 4. *Let c_A be a propositional constraint induced by A over the constraint system $< [0,1], D, V >$. Consider the soft constraint problem $P =< \{c_A\}, con >$, then $c_A \Downarrow con(t)$ is the maximal(best) value of the set $\{\bar{A}(t^{'})|t^{'} \in \Omega_{con_A}, t^{'} \downarrow^{con_A}_{con}= t\}$.*

Corollary 1. *Let c_A be a propositional constraint induced by A over the constraint system $< R^+, +, \times, 0, 1 >$. Consider the soft constraint problem $P =< \{c_A\}, \emptyset >$ then*

$$\tau_n(A) = \frac{1}{n^m} blevel(P).$$

Corollary 2. *Let c_A be a propositional constraint induced by A over the constraint system $< [0,1], D, V >$. Consider the soft constraint problem $P =< \{c_A\}, \emptyset >$, then $blevel(P)$ is the maximal(best) value that $\bar{A}$ can take.*

3.2 A Splitting Algorithm for Computing the Truth Degree of Formulas in Quantitative Logic

The principal question in computing the truth degree of formulas in quantitative logic is to find an optimal method by which this job can be done with less time complexity rather than by directly using the definition.

Lemma 1. *Let c_A, c_B be two propositional constraints induced by A, B over the constraint system $< R^+, +, \times, 0, 1 >$(or$< [0,1], D, V >$). If $con_A \cap con_B \subseteq con \subseteq con_A \cup con_B$, then we have*

$$(c_A \otimes c_B) \Downarrow con = c_A \Downarrow (con \cap con_A) \otimes c_B \Downarrow (con \cap con_B).$$

Lemma 2. *Let c_A be a propositional constraint induced by A over the constraint system $< R^+, +, \times, 0, 1 >$(or$< [0,1], D, V >$). If $con_1 \subseteq con_2 \subseteq con_A$, then we have*
$c_A \Downarrow con_1 = (c_A \Downarrow con_2) \Downarrow con_1$.

Lemma 3. *In the classical propositional logic, suppose A is a formula equivalent to $B \wedge C$. Let c_B, c_C and $c_{B\wedge C}$ be the propositional constraints induced by B, C, and $B \wedge C$ over the constraint system $< R^+, +, \times, 0, 1 >$, then we have $c_{B\wedge C} = c_B \otimes c_C$, i.e., $\forall t \in \Omega_{con_B \cup con_C}$, $c_{B\wedge C}(t) = c_B(t \downarrow con_B) \times c_C(t \downarrow con_C)$.*

Theorem 2. *In the classical propositional logic, suppose A is a formula equivalent to $B \wedge C$. Let c_B, c_C and $c_{B\wedge C}$ be the propositional constraints* *induced by* B, C, *and* $B \wedge C$ *over the constraint system* $< R^+, +, \times, 0, 1 >$. *If* $con_B \cap con_C \subseteq con \subseteq con_B \cup con_C$, *then*

$$\tau_n(A) = n^{-|con_B \cup con_C|} \times \sum_{t \in \Omega_{con}} c_B \Downarrow con \cap con_B(t \downarrow con \cap con_B) \times c_C \Downarrow con \cap con_C(t \downarrow con \cap con_C).$$

Proof. By the Definition 1, Proposition 1, we know

$$\tau_n(A) = n^{-|con_B \cup con_C|} \times \sum_{t' \in \Omega_{con_B \cup con_C}} \overline{B \wedge C}(t').$$

With Lemma 2, it is easy to show that

$$\sum_{t' \in \Omega_{con_B \cup con_C}} \overline{B \wedge C}(t') = \sum_{t \in \Omega_{con}} \sum_{s \in \Omega_{con_B \cup con_C - con}} \overline{B \wedge C}(t, s).$$

According to Definition 5,

$$\sum_{t \in \Omega_{con}} \sum_{s \in \Omega_{con_B \cup con_C - con}} \overline{B \wedge C}(t, s) = \sum_{t \in \Omega_{con}} c_{B \otimes C} \Downarrow con(t).$$

By Lemma 3,

$$\sum_{t \in \Omega_{con}} c_{B \otimes C} \Downarrow con(t) = \sum_{t \in \Omega_{con}} (c_B \otimes c_C) \Downarrow con(t).$$

We learn from Lemma 1 and Definition 5 that

$$\sum_{t \in \Omega_{con}} (c_B \otimes c_C) \Downarrow con(t) = \sum_{t \in \Omega_{con}} (c_B \Downarrow con \cap con_B \otimes c_C \Downarrow con \cap con_C)(t)$$

$$= \sum_{t \in \Omega_{|con|}} c_B \Downarrow con \cap con_B(t \downarrow con \cap con_B) \times c_C \Downarrow con \cap con_C(t \downarrow con \cap con_C).$$

This completes the proof.

Definition 8. *Suppose* $\Gamma = \{A_1, \cdots, A_n\}$, $con \subseteq S$, *then a partition* $\{\Gamma_1, \cdots, \Gamma_m\}$ *of* Γ *is said to be a con-partition if* $\forall i, j, i \neq j$, *we have*

$$d(\otimes \Gamma_i) \cap d(\otimes \Gamma_j) \subseteq t.$$

Definition 9. *Under the assumptions of Definition 8,*

(i)a con-partition $\{\Gamma_1, \cdots, \Gamma_m\}$ *of* Γ *is said to be binary if* $m = 2$;

(ii)a con-partition $\{\Gamma_1, \cdots, \Gamma_m\}$ *of* Γ *is said to be final if* $\forall i = 1, \cdots, m$, *there exists no t-partition of* Γ_i.

Theorem 3. *(splitting algorithm)In the classical propositional logic, suppose A is a formula equivalent to $B = A_1 \wedge A_2 \wedge \cdots \wedge A_m$. Let $\Gamma = \{A_1, \cdots, A_n\}$, $con \subseteq S$, and $\{\Gamma_1, \cdots, \Gamma_m\}$ is a con-partition of Γ. Denote c_{Γ_i} as the propositional constraint induced by $\otimes\Gamma_i$ over the constraint system $< R^+, +, \times, 0, 1 >$,$i = 1, 2, \cdots, m$. Then*

$$\tau_2(A) = 2^{-|con_B|} \times \sum_{t \in \Omega_{con}} (\prod_{i=1}^{m} c_{\Gamma_i} \Downarrow_{con \cap con_{\otimes\Gamma_i}} (t \downarrow^{con}_{con \cap con_{\otimes\Gamma_i}})).$$

Since the t-norm corresponding to the n-valued propositional logic $\prod$ is $\otimes_\pi$, i.e., the product t-norm. We have

Theorem 4. *In the n-valued propositional logic system $\prod$, suppose A is a formula equivalent to $B = A_1 \otimes A_2 \otimes \cdots \otimes A_m$. Let $\Gamma = \{A_1, \cdots, A_n\}$, $con \subseteq S$, and $\{\Gamma_1, \cdots, \Gamma_m\}$ is a con-partition of Γ. Denote c_{Γ_i} as the propositional constraint induced by $\otimes\Gamma_i$ over the constraint system $< R^+, +, \times, 0, 1 >$,$i = 1, 2, \cdots, m$. Then*

$$\tau_n(A) = n^{-|con_B|} \times \sum_{t \in \Omega_{con}} (\prod_{i=1}^{m} c_{\Gamma_i} \Downarrow_{con \cap con_{\otimes\Gamma_i}} (t \downarrow^{con}_{con \cap con_{\otimes\Gamma_i}})).$$

Example 2. Suppose $A = (p_1 \vee \neg p_2 \vee \neg p_5) \wedge (\neg p_1 \vee p_6) \wedge (p_2 \vee \neg p_3) \wedge (p_3 \vee \neg p_4) \wedge (p_4 \vee p_5 \vee \neg p_6)$, let $con = \{p_2, p_4, p_5, p_6\}$, denote the propositional constraints (induced by these five clauses) by c_1, c_2, c_3, c_4, c_5 respectively, then

$$(c_1 \wedge c_2 \wedge c_3 \wedge c_4 \wedge c_5) \Downarrow \{p_2, p_4, p_5, p_6\} = (c_1 \wedge c_2) \Downarrow \{p_2, p_5, p_6\} \times (c_3 \wedge c_4) \Downarrow \{p_2, p_4\} \times c_5.$$

Thus

$$\tau_2(A) = 2^{-6} \times \sum_{(p_2, p_4, p_5, p_6)} N(p_2, p_4, p_5, p_6).$$

where $N(p_2, p_4, p_5, p_6) = c_5(p_4, p_5, p_6) \times \sum_{p_1 \in \{0,1\}} (c_1(p_1, p_2, p_5) \times c_2(p_1, p_6)) \times \sum_{p_3 \in \{0,1\}} (c_3(p_2, p_5, p_3) \times c_4(p_3, p_4))$. At last we get $\tau_2(A) = \frac{12}{64} = 0.1875$.

Remark 1. We should point out that the Theorem 3 does not work in n-valued propositional logic systems such as the L_n, $\mathscr{L}_n^*$, and G_n when $n \geq 3$. The underlying essence is that these t-norms corresponding with these logic systems we have just listed above can't be extended to operations on the set of nonnegative real numbers, and no mention the distributivity of t-norms over the natural addition operation $+$. The Theorem 4 shows us that the n-valued propositional logic system $\prod$ is an exception. This is simply because the corresponding t-norm is the natural product operation $\times$. Obviously $\times$ is distributive over the natural addition operation $+$.

As a matter of fact, in many-valued propositional logic systems, every formula can induce a soft constraint over the soft constraint system $< [0,1], D, V >$. The addition operation in the underlying semirng

$< [0,1], \vee, \otimes, 0, 1 >$ is $\vee$ rather than the natural addition operation on real numbers. Thus this kind of semiring can not be used to count the sum of a certain set of nonnegative numbers. However, the soft constraint system $< [0,1], D, V >$ with the underlying semirng $< [0,1], \vee, \otimes, 0, 1 >$ has its especially advantages. It can be used to compute maximal value that a formula can take semantically. This is an important and interesting question in SCSP problems.

Similarly, in n-valued propositional logic, take the underlying semiring as $< [0,1], D, V >$. We can compute the best value of a given formula as follows.

Theorem 5. *In the n-valued propositional logic, suppose A is a formula equivalent to $B = A_1 \otimes A_2 \otimes \cdots \otimes A_m$. Let $\Gamma = \{A_1, \cdots, A_n\}$, $con \subseteq S$, and $\{\Gamma_1, \cdots, \Gamma_m\}$ is a con-partition of Γ. Denote c_{Γ_i} as the propositional constraint induced by $\otimes\Gamma_i$ over the constraint system $< [0,1], D, V >$,$i = 1, 2, \cdots, m$. Then the best value of the formula A is equal to $blevel(< \{c_B\}, \emptyset >)$ as follows:*

$$blevel(< \{c_B\}, \emptyset >) = \bigvee_{t \in \Omega_{con}} (\otimes_{i=1}^{m} c_{\Gamma_i} \Downarrow_{con \cap con_{\otimes\Gamma_i}} (t \downarrow^{con}_{con \cap con_{\otimes\Gamma_i}})).$$

Define a function $\alpha : [0,1] \to \{0,1\}$ such that $f(a) = 1$ if and only if $a = 1$. Then a propositional constraint c_A induced by formula A can be naturally changed to a constraint $\alpha(c_A)$ over the constraint system $< R^+, D, V >$, where $D = W_n^m, m = |con_A|$. Theorem 3 can be changed to compute the absolute value of formulas in the n-valued propositional logic systems.

Theorem 6. *In the n-valued propositional logic, suppose A is a formula equivalent to $B = A_1 \otimes A_2 \otimes \cdots \otimes A_m$. Let $\Gamma = \{A_1, \cdots, A_n\}$, $con \subseteq S$, and $\{\Gamma_1, \cdots, \Gamma_m\}$ is a con-partition of Γ. Denote c_{Γ_i} as the propositional constraint induced by $\otimes\Gamma_i$ over the constraint system $< [0,1], D, V >$,$i = 1, 2, \cdots, m$. Then the absolute value $\theta_n(A)$ of the formula A is:*

$$\theta_n(A) = n^{-|con_B|} \times \sum_{t \in \Omega_{con}} (\prod_{i=1}^{m} \alpha(c_{\Gamma_i}) \Downarrow_{con \cap con_{A_i}} (t \downarrow^{con}_{con \cap con_{\otimes\Gamma_i}})).$$

4 Conclusion

We should say that our splitting method is not suitable always. For example, in n-valued Łukasiewicz system, our method can not be used to compute the truth degree defined in Definition 1. And even in classical logic, for the kind of formulas $A_1 \wedge \cdots \wedge A_n$ with $d(A_i)$ equal,$i = 1, ..., n$, our method can not be used any longer.

Acknowledgements. This work is supported by National Science Foundation of China (Grant No. 10571112 and 60873119), and the Higher School Doctoral Subject Foundation of Ministry of Education of China under Grant No. 200807180005.

References

1. Zadeh, L.A.: Fuzzy sets. Information Control 8, 338–353 (1965)
2. Pavelka, J.: On fuzzy logic I,II,III. Z Math Logic Grundlagen Math. 25, 45–52,119–134,447–464 (1979)
3. Bistarelli, S.: Semirings for Soft Constraint Solving and Programming. LNCS, vol. 2962. Springer, Heidelberg (2004)
4. Wang, G.J.: Quantitative logic. Chinese Journal of Engineering Mathematics 2, 191–215 (2006)
5. Wang, G.J., Zhou, H.J.: Quantitative logic. Information Sciences 179, 226–247 (2009)
6. Zhou, H.J., Wang, G.J.: A new theory consistency index based on deduction theorems in several logic systems. Fuzzy Sets Syst. 157, 427–443 (2006)
7. Han, B.H., Wang, G.J.: The conditional truth fegree of formulas in classical propositional logic. Fuzzy Systems and Mathematics 4, 9–15 (2007) (in Chinese)
8. Wang, G.J., Hui, X.J.: Randomization of classical inference patterns and its application. Sci. China Ser. E-inf. Sci. 30, 801–812 (2007)
9. Zhang, Z.X., Li, L.F.: Syntactic graded method of two-valued propositional logic formulas. Acta Electronica Sinica 2, 325–330 (2008)
10. Li, J., Zhou, J.S., Zhou, Y.: Absolute truth degree theory of formulas in n-valued Lukasiewicz propositional logic system. Journal of Lanzhou University of Technology 1, 134–1389 (2008) (in Chinese)
11. Wang, G.J., Duan, Q.L.: The (n) truth degree in modal logic and harmonious theorem. Sci. China Ser. F-inf. Sci. 39(2), 234–245 (2009)
12. Han, B.H., Li, Y.M.: Convergency theory in quantitative logic. Computer Engineering and Applications 30, 4–5 (2009) (in Chinese)
13. Han, B.H., Li, Y.M.: Error accumulation theory in quantitative logic. Computer Engineering and Applications 23, 9–10 (2009) (in Chinese)
14. Bayardo, R.J., Pehoushek, J.D.: Counting models using connected components. In: 17th AAAI, pp. 157–162 (2000)
15. Sang, T., Bacchus, F., Beame, P., Kautz, H.A., Pitassi, T.: Combining component caching and clause learning for effective model counting. In: 7th SAT (2004)
16. Darwiche, A.: New advances in compiling CNF into decomposable negation normal form. In: 16th ECAI, pp. 328–332 (2004)
17. Wei, W., Selman, B.: A new approach to model counting. In: Bacchus, F., Walsh, T. (eds.) SAT 2005. LNCS, vol. 3569, pp. 324–339. Springer, Heidelberg (2005)
18. Gomes, C.P., Hoffmann, J., Sabharwal, A., Selman, B.: From sampling to model counting. In: 20th IJCAI, pp. 2293–2299 (2007)
19. Gomes, C.P., Hoffmann, J., Sabharwal, A., Selman, B.: Short XORs for model counting: from theroy to practice. In: Marques-Silva, J., Sakallah, K.A. (eds.) SAT 2007. LNCS, vol. 4501, pp. 100–106. Springer, Heidelberg (2007)
20. Gomes, C.P., Van Hoeve, W.J., Sabharwal, A., Selman, B.: Counting CSP solutions using generalized XOR constraints. In: Proc. of AAAI, pp. 204–209 (2007)

21. Kroc, L., Sabharwal, A., Selman, B.: Leveraging belief propagation, backtrack search, and statistics for model counting. In: Perron, L., Trick, M.A. (eds.) CPAIOR 2008. LNCS, vol. 5015, pp. 127–141. Springer, Heidelberg (2008)
22. Samer, M., Szeider, S.: A fixed-parameter algorithm for ♯ sat with parameter incidence treewidth (2006)
23. Favier, A., de Givry, S., Jegou, P.: Exploiting Problem Structure for Solution Counting. In: Gent, I.P. (ed.) CP 2009. LNCS, vol. 5732, pp. 335–343. Springer, Heidelberg (2009)

Some Remarks on a Few Approximate Reasoning Patterns

Xiao-jing Hui

College of Mathematics and Computer Science,
Yan'an University, Yan'an 716000, P.R. China
e-mail: xhmxiaojing @163.com

Abstract. Three different types of approximate reasoning patterns are proposed in **D**-logic metric space and it is proved that they are equivalent to each other.

Keywords: D- truth degree, **D**-logic metric space, approximate reasoning.

1 Introduction

The probability logic emerged from the 70's of the 20th century (see [1-5]), where uncertainty of premises were considered, and uncertainty degree of the conclusion had been deducted in [1] by using the Kolmogorov axioms[6]. It is remarkable that the theory proposed in [1] is developed by means of individual cases while the probability of one and the same formula varies in different effective inferences, and only a few (mostly two or three) formulas are involved in premises of effective inferences, therefore the theory proposed in [1] seems to be locally but not globally.

On the other hand, a global quantified logic theory is proposed in [7-8] where logic concepts are graded into different levels so as to try to establish a bridge between artificial intelligence and numerical computation (see [9-14]), this can be thought of as continuation and development of the idea of [15-16]. In quantified logic every atomic formula has the same truth degree 0.5, i.e., has the same uncertainty degree 0.5, and any two formulas with the same shape, say, $q_1 \to q_2$ and $q_3 \to q_4$, have the same truth degree. This is not consistent with corresponding problems in the real world. In fact, a simple proposition in the real world is true or not, or in what extent it is true is uncertain. Hence, to follow the way of probabilistic AI and develop a probabilistic style quantified logic is certainly a beneficial task. In view of the above analysis, the paper [17] proposes the concept of **D**-randomized truth degree of formulas by employing a random number sequence, and proves that the set of values of **D**-randomized truth degree of formulas has no isolated point in [0,1]. Moreover, the paper [17] introduces the concept of **D**-similarity degree between formulas, and establishes a **D**- logic metric space without any isolated points, and points out that quantified logic and most of its results can be considered special cases of the new setting. Following above

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 199–206.
springerlink.com

conclusions, the present paper proposed three different types of approximate reasoning patterns in **D**-logic metric space and proved that they are equivalent to each other.

Both truth degree of formulas in quantified logic and **D**-randomized truth degree of formulas in **D**- logic metric space are used to reflect the truth probability of propositions, and hence there is a natural link between the theories of **D-**logic metric spaces and probability logic.

2 Preliminaries

Definition 2.1 [17] *Suppose that* $\mathbf{N} = \{1, 2, \cdots\}$, $\mathbf{D} = (P_1, P_2, \cdots)$, $0 < P_n < 1$ $(n = 1, 2, \cdots)$*, then* ***D*** *is called a random sequence in* $(0, 1)$.

Let $S = \{q_1, q_2, \cdots\}$ be the set of atomic formulas (for distinguishing the symbol of atomic formulas from the symbol of random number $P_n's$, we use $q_n's$ to express atomic formulas), $F(S)$ be the free algebra of type $(\neg, \rightarrow)$ generated by S, the elements of $F(S)$ are formulas.

Definition 2.2[17] *Suppose that* $A = A(q_1, \cdots, q_n) \in F(S)$*, then* A *derives an n-ary Boolean function* $f_A : \{0, 1\}^n \rightarrow \{0, 1\}$ *as follows*

$\forall \alpha = (x_1, \cdots, x_n) \in \{0, 1\}^n$ *, substitutes* x_k *for* q_k *in* $A(q_1, \cdots, q_n)(k = 1, \cdots, n)$ *and keeps the logic connectives* $\neg$ *and* $\rightarrow$ *unchanged and explains them by* $\neg x_k = 1 - x_k$ *and* $x_k \rightarrow x_l = (1 - x_k) \vee x_l$ *respectively, then* $f_A(x_1, \cdots, x_n)$ *is the result of the substitution .* f_A *is said to be induced by* A.

Definition 2.3[17] *Let* $\mathbf{D} = (P_1, P_2, \cdots)$ *be a random sequence in* $(0, 1)$. $\forall \alpha = (x_1, \cdots, x_n) \in \{0, 1\}^n$*, let*

$$\varphi(\alpha) = Q_1 \times \cdots \times Q_n,$$

where $Q_k = P_k$ *when* $x_k = 1$*, and* $Q_k = 1 - P_k$ *when* $x_k = 0 (k = 1, \cdots, n)$*, then we have a mapping*

$$\varphi : \{0, 1\}^n \rightarrow (0, 1),$$

called D-randomized mapping of $\{0,1\}^n$.

It is not difficult to verify that

$$\Sigma\{\varphi(\alpha) \mid \alpha \in \{0,1\}^n\} = 1.$$

Definition 2.4[17] *Suppose that* $A = A(q_1, \cdots, q_n) \in F(S)$, *let*

$$[A] = f_A^{-1}(1), \quad \mu([A]) = \Sigma\{\varphi(\alpha) \mid \alpha \in f_A^{-1}(1)\},$$

and use $\tau_D(A)$ *to denote* $\mu([A])$. *Then call* $\tau_D(A)$ ***D**-randomized truth degree, briefly, **D**-truth degree of* A.

*It is obvious that **D**-randomized truth degree will turn into truth degree proposed in [7] and [10] in case* $P_k = \dfrac{1}{2}$ $(k = 1,2,\cdots)$.

Theorem 2.1[17] *Suppose that* $A \in F(S)$, $\mathbf{D}$ *is a random sequence in (0,1), then* A *is a tautology iff* $\tau_D(A) = 1$, *and* A *is a contradiction iff* $\tau_D(A) = 0$.

Proposition 2.1[17] *Suppose that* $A, B \in F(S)$, *then*

$$\tau_D(A \vee B) = \tau_D(A) + \tau_D(B) - \tau_D(A \wedge B).$$

Definition 2.5[17] *Let **D** be a random sequence in (0,1), let*

$$\xi_D(A,B) = \tau_D((A \to B) \wedge (B \to A)), \qquad A, B \in F(S).$$

Then call $\xi_D(A,B)$ ***D**-similarity degree between* A *and* B.

Theorem 2.2[17] *Let **D** be a random sequence in (0,1), and* $A, B, C \in F(S)$. *Then*

(i) $A \approx B$ iff $\xi_D(A,B) = 1$.

(ii) $\xi_D(A,B) + \xi_D(B,C) \le 1 + \xi_D(A,C)$.

Proposition 2.2[17] *Let **D** be a random sequence in (0,1), let*

$$\rho_D(A,B) = 1 - \xi_D(A,B), \qquad A, B \in F(S),$$

Then ρ_D *is a pseudo-metric on* $F(S)$, *called* **D**-*logic pseudo-metric. Moreover,* $(F(S), \rho_D)$ *is called **D**-logic metric space, which contains no isolated point.*

Remark 1. *Proposition 2.2 shows that* ρ_D *is a reasonable pseudo-metric on* $F(S)$, *because every formula in* $(F(S),\rho_D)$ *can be approximated by a sequence of formulas and this makes it possible to establish an approximate reasoning theory on.*

3 Approximate Reasoning in Randomized Quantified Logic

The three different types of approximate reasoning patterns are proposed in paper [17], but if the three conditions are equivalent to each other ? The question is not answer in paper [17]. The present paper will answer it.

Definition 3.1[17]. *Let* ***D*** *be a random sequence in (0,1) and* Γ *be a logic theory in* ***D****-logic metric space* $(F(S),\rho_D)$, *let*

$$div_D(\Gamma)=\sup\{\rho_D(A,B)\mid A,B\in D(\Gamma)\}.$$

Then $div_D(\Gamma)$ *is said to be the divergent degree of* Γ,Γ *is said to be totally divergent if* $div_D(\Gamma)=1,$ *where* $D(\Gamma)$ *is the set of all* Γ *conclusions.*

Definition 3.2[17]. *Let* ***D*** *be a random sequence in (0,1),* $\Gamma\subset F(S),A\in F(S),\varepsilon>0.$

(i) If $\inf\{\rho_D(A,B)\mid B\in D(\Gamma)\}<\varepsilon,$

then A *is said to be a type-I conclusion of* Γ *with error less than* ε, *and denoted* $A\in D^{\mathbf{1}}_{\varepsilon,D}(\Gamma).$

(ii) If $1-\sup\{\tau_D(B\to A)\mid B\in D(\Gamma)\}<\varepsilon$,

then A *is said to be a type-II conclusion of* Γ *with error less than* ε, *and denoted* $A\in D^{2}_{\varepsilon,D}(\Gamma).$

(iii) If $\inf\{H(D(\Gamma),D(\Sigma))\mid\Sigma\subset F(S),\Sigma\vdash A\}<\varepsilon,$

then A *is said to be a type-III conclusion of* Γ *with error less than* ε, *and denoted* $A\in D^{3}_{\varepsilon,D}(\Gamma),$ *where* H *is the Hausdorff metric on* $\mathsf{P}(F(S)-\{\varnothing\}$.

Proposition 3.1[17] *Let* ***D*** *be a random sequence in* (0,1)*,* $\Gamma\subset F(S),A\in F(S),\varepsilon>0.$ *If* $A\in D^{1}_{\varepsilon,D}(\Gamma).$ *Then* $A\in D^{2}_{\varepsilon,D}(\Gamma).$

Using the following lemma we can proved that if $A\in D_{\varepsilon,}{}^{2}{}_{D}(\Gamma)$, then $A\in D_{\varepsilon,}{}^{1}{}_{D}(\Gamma)$.

Lemma 3.1. *Let **D** be a random sequence in (0,1),* $\Gamma \subset F(S)$, $A \in F(S)$, $B_1 \in D(\Gamma)$, $\alpha \in [0,1]$. *if* $\tau_D(B_1 \to A) = \alpha$, *then there exists* $B_2 \in D(\Gamma)$ *such that* $\tau_D((B_2 \to A) \wedge (A \to B_2)) = \alpha$.

Proof. It followed from $B_1 \in D(\Gamma)$, $\vdash B_1 \to B_1 \vee A$ and the inference rule modus ponens (MP for short) that $B_1 \vee A \in D(\Gamma)$. let $B_2 = B_1 \vee A$. Then

$$\begin{aligned}
&\tau_D((B_2 \to A) \wedge (A \to B_2)) \\
&= \Sigma\{\varphi(\alpha) \mid \alpha \in f^{-1}_{(B_2 \to A)\wedge(A \to B_2)}(1)\}, \\
&= \Sigma\{\varphi(\alpha) \mid \alpha \in f^{-1}_{B_2 \to A}(1) \cap f^{-1}_{A \to B_2}(1)\} \\
&= \Sigma\{\varphi(\alpha) \mid \alpha \in (f^{-1}_{B_2}(1) \cap f^{-1}_{A}(1)) \cup (f^{-1}_{B_2}(0) \cap f^{-1}_{A}(0))\} \\
&= \Sigma\{\varphi(\alpha) \big| f_{B_2}(\alpha) = f_A(\alpha)\} \\
&= \Sigma\{\varphi(\alpha) \mid \alpha \in f^{-1}_{B_1}(0) \cup (f^{-1}_{B_1}(1) \cap f^{-1}_{A}(1))\} \\
&= \tau_D(B_1 \to A) = \alpha
\end{aligned}$$

Theorem 3.1. *Let **D** be a random sequence in (0,1),* $\Gamma \subset F(S)$, $A \in F(S)$, $\varepsilon > 0$, *if* $A \in D_{\varepsilon,\ D}^{\ 2}(\Gamma)$. *Then* $A \in D_{\varepsilon,\ D}^{\ 1}(\Gamma)$.

Proof. Suppose that $A \in D_{\varepsilon,\ D}^{\ 2}(\Gamma)$, then

$$1 - \sup\left\{\tau_D(B \to A) \big| B \in D(\Gamma)\right\} < \varepsilon,$$

it follows from Lemma 3.1, if $B_1 \in D(\Gamma)$ and $\tau_D(B_1 \to A) = \alpha$, then there exists $B_2 \in D(\Gamma)$, such that $\tau_D((B_2 \to A) \wedge (A \to B_2)) = \alpha$. therefore

$$\sup\left\{\tau_D((B_2 \to A) \wedge (A \to B_2)) \big| B_2 \in D(\Gamma)\right\} \geq \sup\left\{\tau_D(B_1 \to A) \big| B_1 \in D(\Gamma)\right\}.$$

It equals to

$$1 - \sup\left\{\tau_D((B_2 \to A) \wedge (A \to B_2)) \big| B_2 \in D(\Gamma)\right\} \leq 1 - \sup\left\{\tau_D(B_1 \to A) \big| B_1 \in D(\Gamma)\right\}.$$

Thus

$$1 - \sup\left\{\tau_D((B_2 \to A) \wedge (A \to B_2)) \big| B_2 \in D(\Gamma)\right\} < \varepsilon,$$

then $\inf\left\{\rho_D(A, B) \big| B \in D(\Gamma)\right\} < \varepsilon$, obviously $A \in D_{\varepsilon,\ D}^{\ 1}(\Gamma)$.

Corollary 3.1. *Let **D** be a random sequence in (0,1),* $\Gamma \subset F(S)$, $A \in F(S)$, $\varepsilon > 0$, $A \in D_{\varepsilon,\ D}^{1}(\Gamma)$ *iff* $A \in D_{\varepsilon,\ D}^{2}(\Gamma)$.

Proposition 3.2[17]. *Let **D** be a random sequence in (0,1),* $\Gamma \subset F(S), A \in F(S), \varepsilon > 0$. *If* $A \in D_{\varepsilon,D}^{3}(\Gamma)$, *then* $A \in D_{\varepsilon,D}^{1}(\Gamma)$.

Theorem 3.2. *Let **D** be a random sequence in (0,1),* $\Gamma \subset F(S)$, $A \in F(S)$, $\varepsilon > 0$, *if* $A \in D_{\varepsilon,\ D}^{1}(\Gamma)$. *Then* $A \in D_{\varepsilon,\ D}^{3}(\Gamma)$.

Proof. (i) Let $\Sigma = \Gamma \cup \{A\}$, $B^* \in D(\Gamma)$, for every $B \in D(\Gamma)$. There exists $A^* \in D(\Sigma)$ (let $A^* = B^*$) satisfies $\rho_D(B^*, A^*) = 0 \le \rho_D(A, B)$. It follows from the definition of metric and infimum that $\rho_D(B^*, D(\Sigma)) \le \rho_D(A, D(\Gamma))$ holds for every $B^* \in D(\Gamma)$. That is to say

$$H_1(D(\Gamma), D(\Sigma)) = \sup\{\rho_D(B^*, D(\Sigma)) \,|\, B^* \in D(\Gamma)\} \le \rho_D(A, D(\Gamma)) < \varepsilon.$$

(ii) Similarly let $\Sigma = \Gamma \cup \{A\}$, $A^* \in D(\Sigma)$, for every $B \in D(\Gamma)$, there exists $\{B_{i_1}, B_{i_2}, \cdots, B_{i_k}\} \subset \Gamma$, such that $\{B_{i_1}, B_{i_2}, \cdots, B_{i_k}\} \vdash B$. There also exists $\{B_{j_1}, B_{j_2}, \cdots, B_{j_m}\} \subset \Gamma$, such that $\{B_{j_1}, B_{j_2}, \cdots, B_{j_m}, A\} \vdash A^*$ .Let $B^* = A^* \vee B_{km}$, where $B_{km} = B_{i_1} \wedge \cdots \wedge B_{i_k} \wedge B_{j_1} \wedge \cdots \wedge B_{j_m}$. It is obviously $B^* \in D(\Gamma)$ and

$$\begin{aligned}
&\tau_D\left(\left(A^* \to B^*\right) \wedge \left(B^* \to A^*\right)\right) \\
&= \Sigma\{\varphi(\alpha) \mid \alpha \in f^{-1}_{(A^* \to B^*) \wedge (B^* \to A^*)}(1)\}, \\
&= \Sigma\{\varphi(\alpha) \mid \alpha \in f^{-1}_{A^* \to B^*}(1) \cap f^{-1}_{B^* \to A^*}(1)\} \\
&= \Sigma\{\varphi(\alpha) \mid \alpha \in (f^{-1}_{A^*}(1) \cap f^{-1}_{B^*}(1)) \cup (f^{-1}_{B^*}(0) \cap f^{-1}_{A^*}(0))\} \\
&= \Sigma\{\varphi(\alpha) \,|\, f_{A^*}(\alpha) = f_{B^*}(\alpha)\}.
\end{aligned}$$

When $f_A(\alpha) = f_B(\alpha) = 0$, it follows from $\vdash B_{km} \to B$ that $f_{B_{km}}(\alpha) = 0$, then $f_{A^*}(\alpha) = f_{B^*}(\alpha)$; when $f_A(\alpha) = f_B(\alpha) = 1$, if $f_{B_{km}}(\alpha) = 0$, then $f_{A^*}(\alpha) = f_{B^*}(\alpha)$;if $f_{B_{km}}(\alpha) = 1$, it follows from $\vdash A \wedge B_{km} \to A^*$, that $f_{A^*}(\alpha)=1$, since $B^* = A^* \vee B_{km}$, hence $f_{A^*}(\alpha)=f_{B^*}(\alpha)$, therefore $\rho_D\left(A^*,B^*\right) \le \rho_D\left(A,B\right)$.

This proves $\rho_D\left(A^*,D\left(\Gamma\right)\right) \le \rho_D\left(A,D\left(\Gamma\right)\right)$ holds for every $A^* \in D\left(\Sigma\right)$, that is to say

$$H_2\left(D\left(\Gamma\right),D\left(\Sigma\right)\right) = \sup\left\{\rho_D\left(A^*,D\left(\Gamma\right)\right)\middle| A^* \in D\left(\Sigma\right)\right\} \le \rho_D(A,D(\Gamma)) < \varepsilon.$$

Following from (i) and (ii) it is obviously that $H\left(D\left(\Gamma\right),D\left(\Sigma\right)\right) < \varepsilon$, This completes the proof.

Corollary 3.2. *Let* ***D*** *be a random sequence in (0,1),* $\Gamma \subset F\left(S\right)$, $A \in F\left(S\right)$, $\varepsilon > 0$, $A \in D_{\varepsilon,\ D}^{\ 1}\left(\Gamma\right)$ *iff* $A \in D_{\varepsilon,\ D}^{\ 3}\left(\Gamma\right)$.

The following conclusion can be obtained from above propositions and theorems:

Conclusion 3.1. *The three conditions given in Definition 3.3 are equivalent to each other.*

3 Conclusion

Based on a random sequence in (0,1), the present paper proposes three different types of approximate reasoning patterns, and proves they are equivalent to each other. more detailed properties will be discussed henceforth.

Acknowledgements. The work was supported by the National Natural Science Foundation of China (No.10771129) and Shaanxi Natural Science Foundation (2010JQ1005).

References

1. Adams, E.W.: A Primer of Probability Logic. CSLI Publications, Stanford (1998)
2. Hailperin, T.: Sentential Probability Logic. Associated University Press, London (1996)
3. Coletti, G., Scozzafava, R.: Probabilistic Logic in a Coherent Setting. Kluwer Academic Publishers, London (2002)
4. Dubois, D., Prade, H.: Possibility theory, probability theory and multiple-valued logics. Annals of Mathematics and Artificial Intelligence 32, 35–66 (2001)
5. Baioletti, M., Capotopti, A., et al.: Simplification rules for the coherent probability assessment problem. Annals of Mathematics and Artificial Intelligence 35, 11–28 (2002)
6. Kolmogorov, N.: The Foundation of Probability. Chelsea Publishing Company, New York (1950)
7. Wang, G.J.: Quantified logic (I). Chinese J. Engineering Math. 32(2), 191–215 (2006) (in Chinese)

8. Wang, G.J.: Introduction to Mathematical Logic and Resolution Principle, 2nd edn. Science in China Press, Beijing (2006) (in Chinese)
9. Wang, G.J., Leung, Y.: Integrated semantics and logic metric space. Fuzzy Sets and Systems 136(1), 71–91 (2003)
10. Wang, G.J., Fu, L., Song, J.S.: Theory of truth degrees of propositions in two-valued propositional logic. Science in China, ser. A 45(9), 1106–1116 (2002)
11. Wang, G.J.: Comparison of deduction theorems in diverse logic systems. New Mathematics and Natural Computation 1(1), 65–77 (2005)
12. Li, B.J., Wang, G.J.: Theory of truth degrees of formulas in Lukasiewicz n-valued propositional logic and a limit theorem. Science in China, ser. E 45(6), 727–736 (2005)
13. Wang, G.J., Qin, X.Y., Zhou, X.N.: Theory of quasi-truth degrees of a kind of formulas in predicate logic. J. Shaanxi Normal University 33(1), 1–6 (2005) (in Chinese)
14. Wang, G.J., Song, J.S.: Graded method in propositional logic (in Chinese). Acta Electronica Sinica 34(2), 252–257 (2006)
15. Pavelka, J.: On fuzzy logic I. II. III. Z. Math. Logic Grundlagen Math. 25, 45–52,119–134,447–464 (1979)
16. Ying, M.S.: A logic for approximate reasoning. Journal of Symbolic Logic 59, 830–837 (1994)
17. Hui, X.J., Wang, G.J.: Randomization of classical inference patterns and its application. Science in China, ser. E 37(6), 801–810 (2007) (in Chinese)

The Interval-Valued Truth Degree Theory of the Modal Formulas

Bi-jing Li[1] and Guo-jun Wang[2]

[1] Department of Mathematics, Baoji College of Arts and Science, Baoji, Shaanxi, 721007, China
e-mail: lbj2007@163.com

[2] College of Mathematics and Information Science , Shaanxi Normal University, Xi'an, Shaanxi 710062, China
e-mail: gjwang@snnu.edu.cn

Abstract. The modal logic, as a decidable fragment of predict logic, not only solved the paradox of material implication thoroughly, but also have important properties. The present paper defines the standard model and the interval-valued truth degree of modal formulas after analyzing the idea of possible world. Then the harmonious theorem is proved, that is, the interval-value truth degree of formulas without modal operators degenerate into a point and the value is just equal to its Borel truth degree.

Keywords: Modal logic, modal operator, Borel probability measure, interval-valued truth degree.

1 Introduction

On the perspective of classical logic, the truth of a proposition is either true or false. Further, we will discover that all propositions of either true or false can be divided into two kinds: one kind is inevitable, and the other is accidental. If we look the necessity and possibility as a relationship between propositions and their truth value, and give this kind of propositions containing above necessity and possibility a new name modal propositions, then the classical proposition logic cannot include the modal propositions, obviously. Consequently, the effectiveness of reasoning between modal propositions can not be reacted. Therefore, we need a new logic system, it is just the modal logic $[1-3]$. In order to formalize the above modal propositions, we employ a new modal operator $[4, 5]$ to analysis the relations between these concepts. Because that modal theory is based on the idea of possible world of Leibniz $[6, 7]$, the truth value of modal propositions is associated with a greater space consisting of possible worlds, as world or time or theory or computing state, etc.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 207–214.
springerlink.com

The modal logic, as a decidable fragment of predict logic, not only solved the paradox of material implication thoroughly, but also have many important properties, for example, satisfiability and strong completeness. Therefore, modal logic has a good prospect of applications $[8-12]$ and much necessity to research.

The reference [13] has already discussed the truth theory of modal formulas, there, it defined n-truth degree using even probability measure on domain, which only contain n different possible world.

2 Preliminaries

In this section, we give a brief introduction on the semantic theory of classical propositional logic $\boldsymbol{L}$, classical modal logic $\boldsymbol{K}$ and the Borel truth degree of a classical formulas without modal operators.

Definition 1. ([14]) *Let* $S = \{p_1, p_2, p_3, \cdots\}$ *be a countable set of atoms or propositional variables. The set of the well formed formulas denoted by* $F(S)$ *is the* $(\neg, \vee, \rightarrow)$*-type free algebra generated by the* S *.*

Definition 2. ([14]) *Classical propositional logic system* $\boldsymbol{L}$ *has three axioms and a inference rule as follows:*

Axioms:

(1) $\varphi \rightarrow (\psi \rightarrow \varphi)$;

(2) $\varphi \rightarrow (\psi \rightarrow \chi) \rightarrow ((\varphi \rightarrow \chi) \rightarrow (\psi \rightarrow \chi))$;

(3) $(\neg\varphi \rightarrow \neg\psi) \rightarrow (\psi \rightarrow \varphi)$;

where $\varphi, \psi, \chi \in F(S)$.

Inference rule: Modus Ponens (shorted by MP).

Definition 3. ([14]) $\boldsymbol{L} - proof$ *is a finite sequence of modal formulas* $\varphi_1, \varphi_2, \cdots, \varphi_n$ *satisfying that* $\forall i \leq n$, φ_i *is either the axiom of the logic system* $\boldsymbol{L}$*, or* $\exists j, k (j, k < i)$*, s.t.* φ_i *is obtained by using MP rule w.r.t.* φ_k *and* φ_j*. In this condition, we call* φ_n *is* $\boldsymbol{L}$*-theorem denoted by* $\vdash_{\boldsymbol{L}} \varphi$.

Definition 4. ([14]) *Let* $B = \{\{0,1\}, \neg, \vee, \rightarrow\}$ *be a Boole algebra, a valuation* v *is a homomorphism from* $F(S)$ *to* $\{0,1\}$*. The set of all valuations is denoted by* Ω*. As a result, we could prove that* $\Omega = 2^\omega = \{0,1\}^\omega$.

Definition 5. ([14]) *Let* $\varphi \in F(S)$. φ *is a tautology if and only if* $\forall v \in \Omega$, $v(\varphi) = 1$.

Definition 6. ([15]) *Let* $X_k = \{0,1\} (k \in N)$ *be the discrete topology spaces, that is,* $\forall k \in N$, $\mathcal{T}_k = \mathcal{P}(\{0,1\}) = \{\emptyset, \{0\}, \{1\}, \{0,1\}\}$*). Suppose* $\Omega = \{0,1\}^\omega = \prod\limits_{k=1}^{\infty} X_k$ *be the product topological space, whose topology* $\mathcal{T}$ *is generated by the topology base* $\mathcal{U}$*, which is family of subsets* $\{A_1 \times \cdots \times A_m \times X_{m+1} \times X_{m+2} \times \cdots \mid A_k \in \mathcal{T}_k, k = 1, 2, \cdots, m, m = 1, 2, \cdots\}$*, then* $(\Omega, \mathcal{T})$ *is called the valuation space.*

Definition 7. ([15]) *A Borel probability measure on valuation space Ω is the probability measure on the $\mathcal{B}(\Omega)$, which is set of all Borel set of topology space Ω. Then we obtain that $\mu(\emptyset) = 0$,$\mu(\Omega) = 1$, and $\mu(\bigcup\limits_{k=1}^{\infty} E_k) = \sum\limits_{k=1}^{\infty} E_k$, where $E_k \in \mathcal{B}(\Omega)$, if $i \neq j$, then $E_i \cup E_j = \emptyset$,$i, j, k = 1, 2, \cdots$.*

Form the reference [16], we could find that $\mathcal{B}(\Omega)$ is the σ-algebra generated by the topology base $\mathcal{U}$.

Definition 8. ([15]) *Let $\varphi \in F(S)$. μ be the Borel probability measure on the valuation space Ω, define that*

$$\tau_\mu(\varphi) = \mu(\varphi^{-1}(1))$$

the $\mu-$probability truth degree of the formula φ.

Definition 9. ([17]) *The language of the classical modal logic* [1, 3] *is generated by the form below:*

$$\varphi := p \mid \perp \mid \neg\varphi \mid \varphi_1 \vee \varphi_2 \mid \diamond\varphi, \quad p \in \Phi$$

where Φ is the set of propositional variables, $\perp$ denotes the contradiction, and
$\Box\varphi = \neg \diamond \neg\varphi$,
$\varphi \wedge \psi = \neg(\neg\varphi \vee \neg\psi)$,
$\varphi \rightarrow \psi = \neg\varphi \vee \psi$.
The set of modal formulas is denoted by Form($\Phi, \diamond$).

Remark 1. In this paper, we want to discuss the truth degree of the modal formulas, here we suppose all of the atoms is countable, i.e., $\Phi \subseteq S$. Under this premise, the modal formulas set $Form(S, \diamond)$ is the biggest one, and $F(S) \subseteq Form(S, \diamond)$. In the present paper, we will construct the theory on the $Form(S, \diamond)$.

Definition 10. ([17]) *A Kripke model for the classical modal logic (or classical model) is a triple $M = (W, R, V)$, where W is a nonempty set of possible worlds, $R \subseteq W \times W$ is a binary relationship on W, and $V : \Phi \rightarrow \mathcal{P}(W)$ is a valuation, where $\mathcal{P}(W)$ is the power set of W.*

Definition 11. ([17]) *Let $\varphi \in Form(\Phi, \diamond)$, $M = (W, R, V)$ be a Kripke model, $w \in W$. The world w satisfying φ, denoted by $M, w \models \varphi$, can be recursively defined as follows:*
(i) $M, w \models p$ if and only if $w \in V(p)$, $p \in \Phi$.
(ii) $M, w \models \perp$ never hold.
(iii) $M, w \models \neg\varphi$ if and only if $M, w \models \varphi$ does not hold.
(iv) $M, w \models \varphi \vee \psi$ if and only if $M, w \models \varphi$ or $M, w \models \psi$.
(v) $M, w \models \diamond\varphi$ if and only if $\exists u \in W$, s.t. $(w, u) \in R$ and $M, u \models \varphi$.
Further, modal formulas φ is valid if and only if $M, w \models \varphi$ holds for every classical model $M = (W, R, V)$ and every world $w \in W$.

Proposition 1. ([17]) *Let $M = (W, R, V)$ be a Kripke model, and define*

$$V(\varphi) = \{w \in W \mid M, w \models \varphi\}, \varphi \in Form(\Phi, \diamond)$$

then

(i) $V(\neg\varphi) = W - V(\varphi)$;
(ii) $V(\varphi \vee \psi) = V(\varphi) \cup V(\psi)$;
(iii) $V(\varphi \wedge \psi) = V(\varphi) \cap V(\psi)$;
(iv) $V(\varphi \to \psi) = (W - V(\varphi)) \cup V(\psi)$;
(v) $V(\diamond\varphi) = \{w \in W \mid R[w] \cap V(\varphi) \neq \emptyset\}$;
where $R[w] = \{u \in W \mid (w, u) \in R\}$.

Definition 12. ([17]) *The classical modal logic system* $\boldsymbol{K}$ *has the following axioms:*

(1) $\varphi \to (\psi \to \varphi)$;
(2) $\varphi \to (\psi \to \chi) \to ((\varphi \to \chi) \to (\psi \to \chi))$;
(3) $(\neg\varphi \to \neg\psi) \to (\psi \to \varphi)$;
(K) $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$.

Modal logic system K has two rules of inference:

(1) MP;
(2) Necessitation : from φ *infer* $\Box\varphi$.

Definition 13. ([17]) $\boldsymbol{K} - proof$ *is a finite sequence of modal formulas* $\varphi_1, \varphi_2, \cdots, \varphi_n$ *satisfying that* $\forall i \leq n$, φ_i *either is the axiom of the system* $\boldsymbol{K}$, *or* $\exists j(j < i)$,*s.t.* φ_i *is the formulas after* φ_j *using necessity rule, or* $\exists j, k(j, k < i)$, *s.t.* φ_i *is obtained by* φ_k *and* φ_j *using MP rule, then we call* φ_n *is* $\boldsymbol{K}$*-theorem, denoted by* $\vdash_{\boldsymbol{K}} \varphi$.

Theorem 1. ([14, 17]) *Classical propositional logic system* $\boldsymbol{L}$ *and classical modal logic* $\boldsymbol{K}$ *are both complete. That is:*

(1) $\forall\varphi \in F(S)$, φ *is the tautology if and only if it is a* $\boldsymbol{L}$*-theorem.*
(2) $\forall\varphi \in Form(S, \diamond)$, φ *is the valid if and only if it is a* $\boldsymbol{K}$*-theorem.*

3 The Interval-Valued Truth Degree

3.1 The Standard Model

In this section, we will analysis the idea of possible world firstly. In order to characterize the truth degree of a formula, at which level a formula becomes true, we will continue the used idea. For a fix formula φ, at the beginning, we must find all of the possible worlds, in which φ is satisfiable. Then we denote this possible worlds set by $V(\varphi)$. Next, we intend to measure this set using measure theory on some probability measure space.

The situation of the set $V(\varphi)$ is influenced by the Kripke model M, $V(\varphi)$ will change if any factor of the M changes. Meanwhile, the concept of the model is basic and flexible in the theory of the modal logic. For example, let

$M_0 = (W, R, V), W_0 = \{w_1, w_2, \cdots\}, V_0(p_1) = \{w_1\}, V_0(p_2) = \{w_2\}$, and $\forall i \in \{3, 4, \cdots\}, V_0(p_i) = \emptyset$. Then we could find that only possible worlds w_1 and w_2 is different to atoms p_1 and p_2 in the sense of satisfiability, and the others are the same w.r.t. all atoms in the sense of satisfiability. Obviously, this model M_0 is inadequate and incomplete, for the reason of laking other various possible worlds, like w_0 which satisfying $\forall i \in N, w_0 \in V(p_i)$. Therefore, the set $V(\varphi)$ under this kind of model can not involve all of the situations in which φ is true, the measure of this $V(\varphi)$ could not performance the truth degree of the φ consequently.

Form above analysis, we need a kind of special model, which involve all possible worlds (i.e., possible situations). On the other hand, modal formulas in modal language are generated by the countable atoms. We look the atoms as not-subdividable constructing factors of possible worlds, then this special model should involve all situations of satisfiability of all atoms.

Definition 14. *Let $M = (W, R, V)$ be a Kripke model. The valuation V induce a function $V^* : W \times S \to \{0, 1\}$ by*

$$V^*(w, p) = 1 \quad \textit{if and only if} \quad w \in V(p),$$

$$V^*(w, p) = 0 \quad \textit{if and only if} \quad w \notin V(p).$$

Definition 15. *Let $M = (W, R, V)$ be a Kripke model. Define a sequence of the binary relations $R_v, R_1, R_2, \cdots$ on W as follows:*

$$wR_V u \quad \textit{if and only if} \quad \forall i \in N,\ V^*(w, p_i) = V^*(u, p_i),$$

$\forall n \in N$

$$wR_n u \quad \textit{if and only if} \quad \forall i \in \{1, 2, \cdots, n\},\ V^*(w, p_i) = V^*(u, p_i).$$

If $wR_V u$, then possible worlds w and u are not diacritical.

Definition 16. *(Standard Model) Let $M = (W, R, V)$ be a model. If*

$$W \cong W/R_V$$

and $|W| = |2^\omega|$, then M is a standard model. We denote the set of all standard models by $\mathcal{M}^*$.

Proposition 2. *Let M be a standard model. Then $M \cong \Omega$, and the isomorphic mapping f works like below:*

$$f(w) = (V^*(w, p_1), V^*(w, p_2), \cdots)$$

Corollary 1. *Let $M_1 = (W_1, R_1, V_1)$ and $M_2 = (W_2, R_2, V_2)$ be both standard models. Then the possible worlds sets are isomorphic, i.e.,*

$$W_1 \cong W_2.$$

There is no difficult to prove that all of the standard models are the same (in the sense of the isomorphism) except the binary relations between the possible worlds.

3.2 The Interval-Valued Truth Degree

Suppose $\Phi = S$, then we could get the relation $F(S) \subseteq Form(\Phi, \diamond)$. Therefore, the truth degree of the modal formulas without modal operators in present paper should be identical to the above Borel truth degree(Definition 8).

Definition 17. *Let $M = (W, R, V)$ be a standard model, W be a product topological space (see Proposition 2), μ be a probability measure on W. Then the Borel probability measure space is $(W, \mathcal{B}(W), \mu)$, where $\mathcal{B}$ is the set of all Borel sets.*

A new problem is following. For a given standard model M, a Borel probability measure μ, a formula φ, the set $V(\varphi)$ is not necessary to be measurable, i.e., under some conditions, $V(\varphi) \notin \mathcal{B}(W)$. The way solving this problem is that we find two sequences of Borel-measurable sets to close the set $V(\varphi)$ form inner and outer respectively.

Definition 18. *Let $M = (W, R, V)$ be a standard model, $A \subseteq W$, $\overline{R}_n(A)$ and $\underline{R}_n(A)$ are upper and lower approximates of A under relation R_n respectively. Then we call $\{\overline{R}_n(A)\}_{n=1}^{\infty}$ and $\{\underline{R}_n(A)\}_{n=1}^{\infty}$ upper and lower approximate sequences of A under model M respectively.*

Proposition 3. *Let M be a standard model, μ be a Borel probability measure on W, $A \subseteq W$. Then*

$$\forall n \in N, \quad \overline{R}_n(A),\ \underline{R}_n(A) \in \mathcal{B}(W),$$

and further

$$\lim_n \mu(\underline{R}_n(A)) \leq \lim_n \mu(\overline{R}_n(A)).$$

Proof. (i) Prove that $\forall n \in N$, $\forall B \subseteq W/R_n$, B is Borel measurable.

Since $\forall n \in N$, W/R_n is finite. The elements W/R_n are all equivalence classes and the closed sets in product topology space W, then every element is Borel measurable. Further, every subset of this finite set W/R_n is Borel measurable.

(ii) Prove that $\forall A \subseteq W$, the sequence $\{\overline{R}_n(A)\}_{n=1}^{\infty}$ is monotone increasing and $\{\underline{R}_n(A)\}_{n=1}^{\infty}$ is monotone decreasing. It is sufficient to prove that $\forall A \subseteq W$, $\forall n \in N$, $\underline{R}_n(A) \subseteq \underline{R}_{n+1}(A)$ and $\overline{R}_{n+1}(A) \subseteq \overline{R}_n(A)$ both hold.

1^0 $\forall w \in \underline{R}_n(A)$, we get $[w]_n \subseteq A$. Because that $[w]_{n+1} \subseteq [w]_n$, then we obtain $[w]_{n+1} \subseteq [w]_n \subseteq A$, i.e., $w \in \underline{R}_{n+1}(A)$, so $\underline{R}_n(A) \subseteq \underline{R}_{n+1}(A)$ holds.

2^0 $\forall w \in \overline{R}_{n+1}(A)$, we get $[w]_{n+1} \cap A \neq \emptyset$. Since $[w]_{n+1} \subseteq [w]_n$, then we can get $[w]_{n+1} \cap A \subseteq [w]_n \cap A$, then $[w]_n \cap A \neq \emptyset$holds, i.e., $w \in \overline{R}_n(A)$, so $\overline{R}_{n+1}(A) \subseteq \overline{R}_n(A)$ holds.

(iii) Form above (i) and (ii), we has proved that the sequence $\mu(\underline{R}_n(A))$ is bounded and monotone increasing and $\mu(\overline{R}_n(A))$ is bounded and monotone decreasing, so the limits of them are both existent. Further, since $\forall n \in N$, $\underline{R}_n(A) \subseteq A \subseteq \overline{R}_n(A)$, we get the conclusion $\lim_n \mu(\underline{R}_n(A)) \leq \lim_n \mu(\overline{R}_n(A))$.

Definition 19. *Let M be a standard model, μ be a Borel probability measure on W, $\varphi \in Form(S, \diamond)$. Define*

$$\tau_{\mu,M}(\varphi) = [\lim_n \mu(\underline{R}_n(A)), \lim_n \mu(\overline{R}_n(A))]$$

the interval-valued truth degree of φ in model M and Borel probability measure μ.

When there is no confusion caused, the lower mark μ and M of $\tau_{\mu,M}(\varphi)$ could be omitted. Denote $\lim_n \mu(\underline{R}_n(A))$ by $\tau_*(\varphi)$ and $\lim_n \mu(\overline{R}_n(A))$ by $\tau^*(\varphi)$ shortly, then $\tau(\varphi) = [\tau_*(\varphi), \tau^*(\varphi)]$.

3.3 The Harmonious Theorem

Theorem 2. *(Harmonious Theorem) Let $\varphi \in F(S)$, $\forall M \in \mathcal{M}^*$, μ be a Borel probability measure on W, then*

$$\tau_{\mu,M}(\varphi) = \tau_\mu(\varphi).$$

Proof. (i) Firstly, we prove that $\forall \varphi \in F(S)$, $\forall M \in \mathcal{M}^*$, there has $\tau_*(\varphi) = \tau^*(\varphi)$, i.e., $\lim_n \mu(\underline{R}_n(A)) = \lim_n \mu(\overline{R}_n(A))$.

Since $\varphi \in F(S)$, φ is a classical propositional formula. Suppose there are k different atoms $p_1, p_2, \cdots, p_k$ included in φ, then $V(\varphi)$ is related with sets $V(p_1), V(p_2), \cdots, V(p_k)$. More detail, $V(\varphi)$ is generated by $V(p_1), V(p_2), \cdots, V(p_k)$ from the operations like intersection, union and complement. Form the Corollary 1, the standard model are isomorphic, then $\forall M \in \mathcal{M}^*$, $V(p_1), V(p_2), \cdots, V(p_k)$ are always the same, consequently, $V(\varphi)$ are the same.

On the other hand, φ consisting finite (k) different atoms, then it is easily to obtain

$$V(\varphi) = \underline{R}_k(V(\varphi)) = \overline{R}_k(V(\varphi)) \in \mathcal{B}(W)$$

and when $l \geq k$, above equation still holds, i.e.,

$$V(\varphi) = \underline{R}_l(V(\varphi)) = \overline{R}_l(V(\varphi)) \in \mathcal{B}(W),$$

then

$$\mu(V(\varphi)) = \mu\underline{R}_l(V(\varphi)) = \mu\overline{R}_l(V(\varphi)), \quad (l \geq k)$$

computing the limits, we get

$$\mu(V(\varphi)) = \tau_*(\varphi) = \tau^*(\varphi).$$

(ii) If we look formulas as mappings $\varphi : W \to \{0, 1\}$, it works like

$$\varphi(w) = 1 \text{ if and only if } w \models \varphi$$

form above (i), we get $\forall M \in \mathcal{M}^*$, $V(\varphi) = \varphi^{-1}(1)$ holds. Therefore,

$$\mu(\varphi^{-1}(1)) = \mu(V(\varphi)) = \tau_*(\varphi) = \tau^*(\varphi),$$

the proof is completed.

Acknowledgements. This work described here is partially supported by the grants from the National Natural Science Foundation of China (NSFC No.10771129).

References

1. Blackburn, P., de Rijke, M., Venema, Y.: Modal Logic. Cambridge University Press, New York (2001)
2. Chagrov, A., Zakharyaschev, M.: Modal Logic. Clarendon Press, Oxford (1997)
3. Chellas, B.F.: Modal Logic: An Introduction. Cambridge University Press, London (1980)
4. Lewis, C.I.: A Survery of Symbolic Logic, 2nd edn. University of California Press, Berkeley (1918/1960)
5. Kripke, S.A.: Semantic considerations on the modal logic. Acta Philos. Fennica 16, 83–94 (1963)
6. Prior, A.N.: Possible Worlds. Philos. Quarterly 12, 36–43 (1962)
7. Mates, B.: The Philosophy of Leibniz. Oxford University Press, Oxford (1986)
8. Prior, A.N.: Time and Modality. Clarendon Press, Oxford (1957)
9. Gochet, P., Gribomont, P.: Handbook of the history of Logic. Epistemic Logic, vol. 7, pp. 99–195. Elsevier, Holland (2006)
10. Koller, D., Pfeffer, A.: Representations and solutions for game-theoretic problems. Artificial Intelligence 4(3), 169–190 (1997)
11. Benalycherif, L., Mclsaac, A.: A Semantic Condition for Data Independence and Application in Hardware Verification. Electronic Notes in Theoretical Computer Science 250(1), 39–54 (2009)
12. Fagin, R., Halpern, J.Y., Moses, Y.: Reasoning About Konwledge. MIT Press, MA (1995)
13. Wang, G.J., Duan, Q.L.: The n-truth Degree Theory of Modal Formulas and Harmonious Theorem. Science in China (Ser. F) 39(2), 234–245 (2009) (in Chinese)
14. Wang, G.J.: An Introduction of Symbolic Logic and Resolution Priciple, 2nd edn. Science Press, Beijing (2008) (in Chinese)
15. Zhong, H.J.: Borel-type Probability Quantitative Logic (submitted for publication) (in Chinese)
16. Cohn, D.L.: Measure Theory. Brikhäuser, Boston (1980)
17. Wang, G.J.: Non-classical Mathematical Logic and Approximate Reasoning, 2nd edn. Science Press, Beijing (2008) (in Chinese)

T-Seminorms and Implications on a Complete Lattice

Hua-wen Liu and Pei-jun Xue

School of Mathematics, Shandong University, Jinan, Shandong 250100, China
e-mail: {hw.liu,pjxue}@sdu.edu.cn

Abstract. An extension of a triangular norm (t-norm for short) called t-seminorm is discussed in this paper. Firstly, we introduce the concept of t-seminorms on a complete lattice. Then, we discuss two kinds of residual implications of t-seminorms, and give the equivalent conditions for infinitely $\vee-$distributive t-seminorms. Furthermore, we define two classes of induced operators of implications on a complete lattice and give the conditions such that they are t-seminorms or infinitely $\wedge-$distributive t-seminorms in their second variables. We also propose another method inducing t-seminorms by implications and another method inducing implications by t-seminorms and involutive negations on a complete lattice.

Keywords: Fuzzy Connective, t-seminorm, Implication, Infinitely $\vee-$distributive, Infinitely $\wedge$-distributive, Closure Operator.

1 Introduction

In fuzzy logics, the set of truth values of fuzzy propositions is modelled by unit interval [0,1] and the truth function for a conjunction connective is usually taken as a triangular norm (t-norm for short) on [0,1] which is monotone, associative, commutative and has neutral element 1 (see [1]). But the t-norms are inadequate to deal with natural interpretations of linguistic words since the axioms of t-norms are quite strong. For instance, when we say " she is very beautiful but stupid", this is not equivalent to "she is very beautiful and stupid". It is in fact " she is very beautiful & stupid" in such a way that & is not a commutative connective but " and " is the common commutative conjunctions (see [2]). In order to interpret the non-commutative conjunctions, Flondor et al. [3] introduced non-commutative t-norms by throwing away the axiom of commutativity of t-norms and used them to construct pseudo-BL-algebras and weak- pseudo-BL-algebras (i.e., pseudo-MTL-algebras [4]). About another axiom of t-norms, i.e. associativity, as underlined in [5,6], for

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 215–225.
springerlink.com

example, "if one works with binary conjunctions and there is no need to extend them for three or more arguments, as happens e.g. in the inference pattern called generalized modus ponens, associativity of the conjunction is an unnecessarily restrictive condition". So, we can obtain another binary truth function for conjunctions by removing the axiom of associativity from the the axioms of non-commutative t-norms, called semi-copula [7,8] (also called t-seminorms [9]). On basis of removing the commutative and associative axioms, Fodor [10,11] proposed weak t-norms on [0,1] and discussed the relations between weak t-norms and implications. Noticing that the QL-implications on [0,1] can not be induced by weak t-norms on [0,1], Wang and Yu [12] generalized the notion of weak t-norms and introduced pseudo-t-norm on a complete Brouwerian lattice L. Further, the relation between the pseudo-t-norms and implications on L was discussed in [12]. Since Wang and Yu's pseudo-t-norms are non-commutative, as the discussion in [3], they should correspond two kinds of residual operators, we call them left and right residual operators respectively. But we find that the left residual operator of a pseudo-t-norm is not an implication in general. So we consider to add conditions $T(x,1) = x(\forall x \in L)$ and that T is nondecreasing in its first variable for Wang and Yu's pseudo-t-norm T, i.e., we slightly strengthen the conditions of Fodor's weak t-norms by replacing the condition $T(x,1) \leq x$ by $T(x,1) = x$ for any $x \in [0,1]$. Thus, we just obtain the definition of t-seminorms in [9].

2 Adjoint Mappings on a Complete Lattice

In this section, we briefly recall the definition of adjoint mappings and discuss some of the properties for our usage.

Definition 2.1.([13]) *Let X and Y be two posets, the mappings $f\colon X \to Y$ and $g : Y \to X$ be non-decreasing. We call f is a left adjoint mapping of g, or g is a right adjoint mapping of f, and write $f \dashv g$, if the following adjuntion condition holds for all $x \in X$ and $y \in Y$:*

$$f(x) \leq y \text{ if and only if } x \leq g(y)$$

Proposition 2.1.([13]) *Let X and Y be posets, the mappings $f : X \to Y$ and $g : Y \to X$ be non-decreasing.*

(i) $f \dashv g$ if and only if $x \leq g \circ f(x)$, $f \circ g(x) \leq y$ for all $x \in X, y \in Y$.

(ii) if $f \dashv g$, then $f \circ g \circ f = f$ and $g \circ f \circ g = g$.

Proposition 2.2.([13]) *Suppose that X and Y are complete lattices, then the following properties hold:*

(i) An non-decreasing $f\colon X \to Y$ has a right adjoint mapping if and only if f is infinitely $\vee-$distributive, i.e., $f(\sup_{z\in Z} z) = \sup_{z\in Z} f(z)$, where Z is any nonempty subset of X.

(ii) An non-decreasing $g\colon Y \to X$ has a left adjoint mapping if and only if g is infinitely $\wedge-$distributive, i.e., $g(\inf_{z\in Z} z) = \inf_{z\in Z} g(z)$, where Z is any nonempty subset of Y.

Definition 2.2.([13]) *Let X be a poset and mapping $t : X \to X$ non-decreasing. We call t a closure operator if the following hold:*
(i) $x \le t(x)$ for all $x \in X$;
(ii) $t \circ t = t$.

Definition 2.3.([14]) *Let X be a complete lattice. (L, t) is called an inference system if $t : X \to X$ is a closure operator.*

Proposition 2.3.([13]) *Let X be a poset and mapping $t : X \to X$ non-decreasing. Then t is a closure operator if and only if there exists a poset Y and two non-decreasing mappings $f\colon X \to Y$ and $g\colon Y \to X$ such that $f \dashv g$ and $t = g \circ f$.*

From Proposition 2.3, we know that if we take $X = Y = L$ (a complete lattice) and for any fixed $a \in L$, $f(x) = T(a, x)$, $g(x) = I(a, x)$ for any $x \in L$, where T is a t-norm on L such that (T, I) forms an adjoint couple on L, then $t : L \to L$, i.e., $t(x) = I(a, T(a, x))$ for any $x \in L$ is a closure operator, and hence (L, t) forms an inference system.

Definition 2.4. *Let X and Y be complete lattices, $h : X \to Y$. We define $f_h : Y \to X$ and $g_h : Y \to X$ as follows:*

$$f_h(y) = \inf\{t \in X \mid y \le h(t)\}, \quad \forall y \in Y. \tag{2.1}$$

$$g_h(y) = \sup\{t \in X \mid h(t) \le y\}, \quad \forall y \in Y. \tag{2.2}$$

We can easily obtain the following results from Proposition 2.3.

Theorem 2.1. *Suppose that X and Y are two complete lattices, $h : X \to Y$ is a non-decreasing mapping and f_h, g_h are defined by (2.1) and (2.2) respectively.*
(i) *h is infinitely $\vee-$distributive if and only if $h \dashv g_h$;*
(ii) *h is infinitely $\wedge-$distributive if and only if $f_h \dashv h$.*

Theorem 2.2. *Suppose that X and Y are two complete lattices, $h : X \to Y$ is a non-decreasing mapping and f_h, g_h are defined by (2.1) and (2.2) respectvely.*
(i) h is infinitely $\vee-$distributive if and only if $g_h(y) =$ $\max\{t \in X \mid h(t) \le y\}, \quad \forall y \in Y$;
(ii) h is infinitely $\wedge-$distributive if and only if $f_h(y) =$ $\min\{t \in X \mid y \le h(t)\}, \quad \forall y \in Y$.

Proof. We use the same notation $\leq$ to denote the partial orders in X and Y.

(i) If h is infinitely $\vee-$distributive, then it follows from Theorem 2.1 that $h \dashv g_h$. So, for any $y \in Y$, $g_h(y) \leq g_h(y)$ follows that $h(g_h(y)) \leq y$, i.e., $g_h(y) \in \{t \in X \mid h(t) \leq y\}$. Therefore, $g_h(y) = \max\{t \in X \mid h(t) \leq y\}$.

Conversely, suppose that $g_h(y) = \max\{t \in X \mid h(t) \leq y\}$ holds for any $y \in Y$. We need to prove $h(\sup\limits_{z \in Z} z) = \sup\limits_{z \in Z} h(z)$ for any nonempty subset Z of X.

Observe firstly that from the monotonicity of h ,we always have $h(\sup\limits_{z \in Z} z) \geq \sup\limits_{z \in Z} h(z)$.

Let $b = h(\sup\limits_{z \in Z} z)$. This implies that $h(z) \leq b$ for any $z \in Z$. Therefore, $z \in \{t \in X \mid h(t) \leq b\}$ for every $z \in Z$, and consequently $z \leq g_h(b)$ for every $z \in Z$. Thus, $\sup\limits_{z \in Z} z \leq g_h(b)$. From the monotonicity of h, we get $h(\sup\limits_{z \in Z} z) \leq h(g_h(b)) \leq b = \sup\limits_{z \in Z} h(z)$. So we obtain $h(\sup\limits_{z \in Z} z) = \sup\limits_{z \in Z} h(z)$.

(ii) Similar to the proof of (i).

The following are the straight conclusions of Proposition 2.3 and Theorem 2.1.

Theorem 2.3. *Let X and Y be two complete lattices. Suppose that $h : X \to Y$ is a non-decreasing and f_h, g_h are defined by (2.1) and (2.2) respectively.*

(i) *If h is infinitely $\vee-$distributive, then $t_1 = g_h \circ h$ is a closure operator on X and hence (X, t_1) forms an inference system.*

(ii) *If h is infinitely $\wedge-$distributive, then $t_2 = h \circ f_h$ is a closure operator on Y and hence (Y, t_2) forms an inference system.*

3 T-Seminorms and Their Residual Implications

In the following, we always use L to denote a complete lattice with the maximal element 1 and minimal element 0.

Definition 3.1. *A binary operation I on L is an implication if it satisfies:*

(i) $I(1, y) = y$ and $I(0, y) = 1$ for any $y \in L$;

(ii) I is non-increasing in its first and non-decreasing in its second variable.

Remark 3.1. *The above definition of implications on L is different from the one in [13], where there is no the condition: I is non-increasing in its first variable.*

Form Definition 3.1, we know $I(x, 1) = 1$ for any $x \in L$ since $I(x, 1) \geq I(1, 1) = 1$.

Definition 3.2. (Fodor [10,11]) *A function $f : [0, 1]^2 \longrightarrow [0, 1]$ is called a weak t-norm if it satisfies the following conditions:*

(i) $T(x,1) \leq x$, $T(1,y) = y$ *for any* $x, y \in [0,1]$*;*

(ii) $T(x_1, y_1) \leq T(x_2, y_2)$ *if* $x_1 \leq x_2$ *and* $y_1 \leq y_2$ *for any* $x_1, x_2, y_1, y_2 \in [0,1]$.

Wang and Yu [12] point out that Fodor's definition of weak t-norms on [0,1] can be easily extended to a general complete Brouwerian lattice. Noticing that general QL–implication on [0,1] cannot be induced by weak t-norms on [0,1] (see [10]), Wang and Yu [12] generalized the notion of weak t-norms to the following form.

Definition 3.3. (Wang and Yu[12]) *A binary operation* T *on* L *is called a pseudo-t-norm if it satisfies the following conditions:*

(i) $T(1,x) = x$, $T(0,y) = 0$ *for any* $x, y \in L$*;*

(ii) $y \leq z$ *implies* $T(x,y) \leq T(x,z)$ *for all* $x, y, z \in L$.

Remark 3.2. *There are two kinds of residual operations for a pseudo-t-norm or for a weak t-norm, but one of them is not an implication on* L *in general. In fact, let* T *be a pseudo-t-norm or a weak t-norm on* L*, we cannot determine the value of* $\sup\{t \in L \mid T(t,1) \leq y\}$ *since* $T(t,1)$ *cannot be determined for a given* $t \in L$*. This means that for the residual operator* I_T *of* T*, we cannot judge if* $I_T(1,y) = y$ *holds for any* $y \in L$*. This fact impels us to generalize the notions of pseudo-t-norms and weak t-norms as follows.*

Definition 3.4. *A binary operation* T *on* L *is called a triangular seminorm (briefly t-seminorm) if it satisfies the following conditions:*

(i) $T(1,x) = T(x,1) = x$ *for all* $x \in L$*;*

(ii) T *is non-decreasing in each variable.*

It is clear from above definition that $T(x,0) = T(0,x) = 0$ and $T_D(x,y) \leq T(x,y) \leq T_M(x,y)$ for any $x, y \in L$, where $T_D(x,y) = min(x,y)$ if $max(x,y) = 1, T_D(x,y) = 0$ otherwise, and $T_M(x,y) = min(x,y)$ ([15]). Moreover, any t-seminorm on L must be a weak t-norm and hence a pseudo-t-norm on L. A t-seminorm on L is a t-norm only when it is commutative and associative.

Definition 3.5. *Let* T *be a t-seminorm on* L*. The following* I_{1T} *and* I_{2T}*:* $L^2 \to L$ *are said to be type-1 and type-2 residual operators respectively, for any* $x, y \in L$*,*

$$I_{1T}(x,y) = \sup\{t \in L \mid T(t,x) \leq y\}, \tag{3.1}$$

$$I_{2T}(x,y) = \sup\{t \in L \mid T(x,t) \leq y\}. \tag{3.2}$$

Obviously, $I_{1T} = I_{2T}$ when t-seminorm T is commutative. It is easy to verify that type-1 and type-2 residual operators of t-seminorm T on L are implications on L. Moreover, if T_1 and T_2 are comparable t-seminorms on L such that $T_1 \leq T_2$, then $I_{1T_1} \geq I_{1T_2}$ and $I_{2T_1} \geq I_{2T_2}$.

For a given $y \in L$, if we write $h(x) = T(y, x)$ for all $x \in L$, then h is a non-decreasing mapping on L and I_{2T} corresponds to formula (2.2). The following are the straight results of Proposition 2.1, Theorems 2.1 and 2.2.

Theorem 3.1. *Let T be a t-seminorm on L. Then the following statements are equivalent:*

(i) T is infinitely $\vee-$distributive in its second variable;

(ii) For any fixed $x \in L$, $T(x, \cdot) \dashv I_{2T}(x, \cdot)$, i.e., $T(x, y) \le z$ if and only if $y \le I_{2T}(x, z)$ for any $y, z \in L$;

(iii) $y \le I_{2T}(x, T(x, y))$, $T(x, I_{2T}(x, y)) \le y$ for all $x, y \in L$;

(iv) $I_{2T}(x, y) = \max\{t \in L \mid T(x, t) \le y\}$ for any $x, y \in L$.

The equivalentness about (i), (ii) and (iv) in Theorem 3.1 is same as the result for pseudo-t-norms given by Wang and Yu in [12].

From Proposition 2.2, Theorems 2.3 and 3.1, we can get the following results.

Theorem 3.2. *Let T be a t-seminorm on L and satisfy infinitely $\vee-$distributive law in its second variable. Then*

(i) $T(x, I_{2T}(x, T(x, y))) = T(x, y)$, $I_{2T}(x, T(x, I_{2T}(x, y))) = I_{2T}(x, y)$ for all $x, y \in L$.

(ii) For any given $y \in L$, the mapping $t : L \to L$ defined by $t(x) = I_{2T}(y, T(y, x))$ is a closure operator on L , and hence (L, t) forms an inference system.

Theorem 3.2 (i) is Wang and Yu's result for pseudo-t-norm T (see Theorem 4.3 in [13]).

For any t-seminorm T on L, if we define $T' : L^2 \to L$ by $T'(x, y) = T(y, x)$ for any $x, y \in L$, then it is clear that T' is also a t-seminorm on L. So we know $I_{1T} = I_{2T'}$. Therefore, we can get from Theorems 3.1 and 3.2 the following corresponding results associated to the type-1 residual implications.

Theorem 3.3. *Let T be a t-seminorm on L. Then the following statements are equivalent:*

(i) T is infinitely $\vee-$distributive in its first variable;

(ii) For any given $y \in L$, $T(\cdot, y) \dashv I_{1T}(y, \cdot)$, i.e., $T(x, y) \le z$ if and only if $x \le I_{1T}(y, z)$ for all $x, z \in L$;

(iii) $y \le I_{1T}(x, T(y, x))$, $T(I_{1T}(x, y), x) \le y$ for any $x, y \in L$;

(iv) $I_{1T}(x, y) = max\{t \in L \mid T(t, x) \le y\}$ for any $x, y \in L$.

Theorem 3.4. *Let T be a t-seminorm on L and satisfy infinitely $\vee-$distributive law in its first variable. Then*

(i) $T(I_{1T}(y, T(x, y)), y) = T(x, y)$, $I_{1T}(x, T(I_{1T}(x, y), x)) = I_{1T}(x, y)$ for any $x, y \in L$;

(ii) For any given $y \in L$, the mapping $t : L \to L$ defined by $t(x) = I_{1T}(y, T(x, y))$ is a closure operator on L , and hence (L, t) forms an inference system.

Theorem 3.5. *Let T be a t-seminorm on L and infinitely $\vee-$distributive in its two variables. Then I_{1T} and I_{2T} satisfy*

(OP) $I_{iT}(x,y) = 1$ if and only if $x \leq y$ for all $x, y \in L, i = 1, 2$;

($I_\wedge$) I_{iT} is infinitely $\wedge-$distributive in its second variable, $i = 1, 2$.

Proof. (OP) It follows from Theorem 3.3 that for any $x, y \in L$, $x \leq y$ if and only if $T(1, x) \leq y$ if and only if $1 \leq I_{1T}(x, y)$ if and only if $I_{1T}(x, y) = 1$ The proof for the case of I_{2T} is similar.

($I_\wedge$) For any $x \in L$, and any subset Y of L, if $Y = \emptyset$, then $I_{1T}(x, \inf_{y \in Y} y) = I_{1T}(x, 1) = 1 = \inf_{y \in Y} I_{1T}(x, y)$ since $\inf \emptyset = 1$; if $Y \neq \emptyset$, then we have

$$
\begin{aligned}
&I_{1T}(x, \inf_{y \in Y} y) = \sup\{t \in L \mid T(t, x) \leq \inf_{y \in Y} y\} \\
&= \sup\{t \in L \mid \forall y \in Y, T(t, x) \leq y\} = \sup\{t \in L \mid \forall y \in Y, t \leq I_{1T}(x, y)\} \\
&= \sup\{t \in L \mid t \leq \inf_{y \in Y} I_{1T}(x, y)\} = \inf_{y \in Y} I_{1T}(x, y).
\end{aligned}
$$

The proof for the case of I_{2T} is similar.

4 T-Seminorms Induced by Implications on a Complete Lattice

In this section, we discuss the t-seminorms induced by implications on a complete lattice.

Definition 4.1. *Let $I : L^2 \to L$ be an implication. We define the induced operators T_{1I} and T_{2I} from I as follows, for any $x, y \in L$,*

$$T_{1I}(x, y) = \inf\{t \in L \mid x \leq I(y, t)\}. \tag{4.1}$$

$$T_{2I}(x, y) = \inf\{t \in L \mid y \leq I(x, t)\}. \tag{4.2}$$

Obviously, $T_{1I} = T_{2I}$ holds if I satisfies $x \leq I(y, z)$ iff $y \leq I(x, z)$ for any $x, y, z \in L$.

Remark 4.1. *(i) The T_{1I} and T_{2I} defined by (4.1) and (4.2) are two well defined operators, i.e., the appropriate sets in (4.1) and (4.2) are non-empty since $I(x, 1) = 1$ for any $x \in L$.*

(ii) It is worthwhile to mention that T_{1I} and T_{2I} defined by the above are not necessarily t-seminorms. For instance, if we take $I(x, y) = 1 - x + xy$ for all $x, y \in [0, 1]$ (Reichenbach implication), then for any $x > 0$, by (4.1) and (4.2), we obtain that $T_{1I}(1, x) = 1 \neq x$ and $T_{2I}(x, 1) = 1 \neq x$. These facts mean that I_{1T} and I_{2T} are not t-seminorms.

We now give the conditions such that T_{1I} and T_{2I} are t-seminorms.

Theorem 4.1. *Assume that I is an implication on L and satisfies (OP): $I(x,y)=1$ if and only if $x \le y$ for any $x,y \in L$. Then T_{1I} and T_{2I} defined by (4.1) and (4.2) are t-seminorms on L .*

Proof. Since I is an implication on L and satisfies (OP), we get by (4.1) that for any $x \in L$,

$$T_{1I}(1,x) = \inf\{t \in L \mid 1 \le I(x,t)\} = \inf\{t \in L \mid x \le t\} = x,$$

$$T_{1I}(x,1)\inf\{t \in L \mid x \le I(1,t)\} = \inf\{t \in L \mid x \le t\} = x.$$

For any $x_1, x_2, y_1, y_2 \in L$ and $x_1 \le x_2, y_1 \le y_2$, since $y_1 \le y_2$ implies $I(y_2,t) \le I(y_1,t)$ for any $t \in L$, we have $x_1 \le x_2 \le I(y_2,t) \le I(y_1,t)$, i.e., $t \in \{t \in L \mid x_1 \le I(y_1,t)\}$, if $t \in \{t \in L \mid x_2 \le I(y_2,t)\}$. This means that $\{t \in L \mid x_2 \le I(y_2,t)\} \subseteq \{t \in L \mid x_1 \le I(y_1,t)\}$. So we get $\inf\{t \in L \mid x_1 \le I(y_1,t)\} \le \ inf\{t \in L \mid x_2 \le I(y_2,t)\}$, i.e., $T_{1I}(x_1,y_1) \le T_{1I}(x_2,y_2)$ when $x_1 \le x_2$ and $y_1 \le y_2$.

Therefore, T_{1I} is a t-seminorm. The proof for T_{2I} is similar to the above.

For any fixed $x \in L$, by taking $h(y) = I(x,y)$ for all $y \in L$, we obtain the following straight results of Proposition 2.1 and Theorems 2.1 and 2.2.

Theorem 4.2. *Let I be an implication on L. Then the following statements are equivalent:*

($I_\wedge$) I is infinitely $\wedge-$distributive in its second variable;

(i) for any fixed $y \in L$, $T_{1I}(\cdot,y) \dashv I(y,\cdot)$;

(ii) $x \le I(y, T_{1I}(x,y))$, $T_{1I}(I(y,x),y) \le x$ for any $x,y \in L$;

(iii) $T_{1I}(x,y) = \min\{t \in L \mid x \le I(y,t)\}$ for any $x,y \in L$.

Theorem 4.3. *Let I be an implication on L. Then the following statements are equivalent:*

($I_\wedge$) I is infinitely $\wedge-$distributive in its second variable;

(i) For any fixed $x \in L$, $T_{2I}(x,\cdot) \dashv I(x,\cdot)$;

(ii) $y \le I(x, T_{2I}(x,y))$, $T_{2I}(x, I(x,y)) \le y$ for any $x,y \in L$;

(iii) $T_{2I}(x,y) = \min\{t \in L \mid y \le I(x,t)\}$ for any $x,y \in L$.

From Proposition 2.2 and Theorems 4.2, 4.3 and 2.3 we can get the following results.

Theorem 4.4. *Let I be an implication on L satisfying infinitely $\wedge-$distributive law in its second variable. Then*

(i) $T_{1I}(I(y,T_{1I}(x,y)),y) = T_{1I}(x,y)$, $I(y,T_{1I}(I(y,x),y)) = I(y,x)$ for any $x,y \in L$;

(ii) $T_{2I}(x,I(x,T_{2I}(x,y))) = T_{2I}(x,y)$, $I(x,T_{2I}(x,I(x,y))) = I(x,y)$ for any $x,y \in L$;

(iii) For any fixed $y \in L$, the mapping $t_1, t_2 : L \to L$ defined by $t_1(x) = I(y, T_{1I}(x,y))$ and $t_2(x) = I(y, T_{2I}(y,x))$ for all $x \in L$ are closure operators on L and hence (L,t_1) and (L,t_1) form two inference systems.

We now give the conditons such that T_{1I} and T_{2I} are t-seminorms and they are infinitely $\vee-$distributive respectively in its first and in its second variable.

Theorem 4.5. *Let I be an implication on L and satisfy (OP) and ($I_\wedge$). Then T_{1I} defined by (4.1) and T_{2I} defined by (4.2) are t-seminorms satisfying infinitely $\vee-$distributive law respectively in its first and in its second variable. Moreover, $I = I_{1T_{1I}} = I_{2T_{2I}}$.*

Proof. First, it follows from Theorem 4.1 that T_{1I} and T_{2I} are t-seminorms on L under the assumptions. We now prove that T_{1I} and T_{2I} are infinitely $\vee-$distributive respectively in its first and in its second variable.

From Theorem 4.2 we know $T_{1I}(\cdot, y) \dashv I(y, \cdot)$ for any fixed $y \in L$. Since implication I satisfies $\mathrm{I}_\wedge$, so we have, for any $y \in L$ and any subset X of L,

$$\begin{aligned} T_{1I}(\sup_{x\in X} x, y) &= \inf\{t \in L \mid \sup_{x\in X} x \le I(y,t)\} \\ &= \inf\{t \in L \mid \forall x \in X, x \le I(y,t)\} \\ &= \inf\{t \in L \mid \forall x \in X, T_{1I}(x,y) \le t\} \\ &= \inf\{t \in L \mid \sup_{x\in X} T_{1I}(x,y) \le t\} \\ &= \sup_{x\in X} T_{1I}(x,y). \end{aligned}$$

The proof for the case of T_{2I} is similar.

In the sequel, we consider $I_{1T_{1I}}$ and $I_{2T_{2I}}$. By means of Theorems 3.1 and 3.3, we have, for any $x, y \in L$,

$$I_{1T_{1I}}(x,y) = \sup\{t \in L \mid T_{1I}(t,x) \le y\} = \sup\{t \in L \mid t \le I(x,y)\} = I(x,y)$$

$$I_{2T_{2I}}(x,y) = \sup\{t \in L \mid T_{2I}(x,t) \le y\} = \sup\{t \in L \mid t \le I(x,y)\} = I(x,y)$$

So we obtain $I_{1T_{1I}} = I_{2T_{2I}} = I$ if implication I satisfies (OP) and ($\mathrm{I}_\wedge$).

Summarize the results in Theorems 3.5 and 4.5, we get the following theorem.

Theorem 4.6. *Let T be a t-seminorm on L.*

(i) If T is infinitely $\vee-$distributive in its first variable, then I_{1T} satisfies (OP) and ($I_\wedge$), and $T = T_{1I_{1T}}$. Conversely, if implication I on L satisfies (OP) and ($I_\wedge$), then T_{1I} is infinitely $\vee-$distributive in its first variable, and $I = I_{1T_{1I}}$.

(ii) If T is infinitely $\vee-$distributive in its second variable, then I_{2T} satisfies (OP) and ($I_\wedge$), and $T = T_{2I_{2T}}$. Conversely, if implication I on L satisfy (OP) and ($I_\wedge$), then T_{2I} is infinitely $\vee-$distributive in its second variable, and $I = I_{2T_{2I}}$.

Proof. We only need to prove $T = T_{1I_{1T}}$ in (i) and $T = T_{2I_{2T}}$ in (ii), since the others are all the results in Theorems 3.5 and 4.5. We only prove the first equation since the proof for another is similar.

It follows from Theorems 3.3 that for any $x, y \in L$,

$$\begin{aligned} T_{1I_{1T}}(x,y) &= \inf\{t \in L \mid x \le I_{1T}(y,t)\} \\ &= \inf\{t \in L \mid T(x,y) \le t\} = T(x,y). \end{aligned}$$

We now propose another method inducing t-seminorms from implications on L and another method inducing implications from t-seminorms and involuative negations on L.

Theorem 4.7. *(i) Let I be an implication on L and the negation $N_I : L \to L$ defined by: $N_I(x) = I(x, 0)$ for any $x \in L$ is involuative. Then the mapping $T : L^2 \to L$ defined by*

$$T(x, y) = N_I(I(x, N_I(y))),\ x, y \in L \tag{4.3}$$

is a t-seminorm on L.

(ii) Let T be a t-seminorm on L and N an involuative negation on L. Then $I : L^2 \longrightarrow L$ defined by

$$I(x, y) = N(T(x, N(y))),\ x, y \in L \tag{4.4}$$

is an implication on L.

Proof (i) First of all, it follows from (4.3) we know that T is non-decreasing in both variables. For any $x \in L$, by (4.3) we get

$$T(x, 1) = N_I(I(x, N_I(1))) = N_I(I(x, 0)) = x,$$

$$T(1, x) = N_I(I(1, N_I(x))) = N_I(N_I(x)) = x.$$

(ii) From (4.4) we know that I is non-increasing in its first variable and non-decreasing in its second variable. For any $y \in L$, we have by (4.4) that

$$I(1, y) = N(T(1, N(y))) = N(N(y)) = y.$$

$$I(0, y) = N(T(0, N(y))) = N(0)) = 1.$$

5 Conclusion

In this paper, we have introduced the definition of t-seminorms on a complete lattice, and defined two types of residual operators of t-seminorms. We have pointed out that the residual operators of t-seminorms are all implications on a complete lattice and their properties have been discussed. The equivalent conditions of t-seminorms satisfying infinitely $\vee-$distributive law respectively in its first and in its second variable have been given. For an implication I on a complete lattice, we have defined its two kinds of induced operators T_{1I} and T_{2I} and given the conditions such that T_{1I} and T_{2I} are t-seminorms and the conditions such that $T_{1I}(T_{2I})$ is a t-seminorm satisfying infinitely $\vee-$distributive law in its first (second) variable. The equivalent conditions for the implications satisfying infinitely $\wedge-$distributive law in its second variable have also been given. We have also proposed another method inducing t-seminorms from implications and another method inducing implications from t-seminorms and involutive negations on a complete lattice. In our future study, we will generalize this work to the case of uninorms on a complete lattice. We will define semi-uninorms and further define and discuss their residual operators. These works will bring benefit for approximate reasoning, information aggregation and other application areas.

Acknowledgements. This work described here is supported by the grants from the National Natural Science Foundation of China (NSFC No.60774100) and the Natural Science Foundation of Shandong Province of China (No.Y2007A15).

References

1. Klement, E.P., Mesiar, R., Pap, E.: Triangular Norms. Kluwer Academic Publishers, Dordrecht (2000)
2. Eslami, E., Khosravi, H., Sadeghi, F.: Very and more or less in non-commutative fuzzy logic. Soft-Computing 12, 275–279 (2008)
3. Flondor, P., Georgescu, G., Lorgulescu, A.: Pseudo-t-norms and pseudo-BL algebras. Soft Computing 5, 355–371 (2001)
4. Hájek, P.: Observation on non-commutative fuzzy logic. Soft-Computing 8, 38–43 (2003)
5. Fodor, J., Keresztfalvi, T.: Nonstandard conjuctions and implications in fuzzy logic. International Journal of Approximate Reasoning 12, 69–84 (1995)
6. Kawaguchi, M.F., Miyakoshi, M.: Composite fuzzy relational equations with non-commutative conjunctions. Information Sciences 110, 113–125 (1998)
7. Bassan, B., Spizzichino, F.: Relations among univariate aging bivariate aging and dependence for exchangeable lifetimes. J. Multivariate Anal. 93, 313–339 (2005)
8. Durante, F., Klement, E.P., Mesiar, R., et al.: Conjunctors and their R esidual Implicators: Characterizations and Construct Methods. Mediterranean Journal of Mathematics 4, 343–356 (2007)
9. Suárez Garcia, F., Gil Álvarez, P.: Two families of fuzzy integrals. Fuzzy Sets and Systems 18, 67–81 (1986)
10. Fodor, J.: On fuzzy implication operators. Fuzzy Sets and Systems 42, 293–300 (1991)
11. Fodor, J.: Strict preference relations based on weak t-norms. Fuzzy Sets and Systems 43, 327–336 (1991)
12. Wang, Z., Yu, Y.: Pseudo-t-norms and implication operators on a complete Brouwerian lattice. Fuzzy Sets and Systems 132, 113–124 (2002)
13. Zheng, C., Fan, L., Cui, H.: Frame and Continuous Lattice, 2nd edn. Capital Normal University Press, Beijing (1994)
14. Wang, G.J.: Non-classical Mathematical Logic and Approximate Reasoning, 2nd edn., pp. 144–146. Science Press, Beijing (2003)
15. Durante, F., Mesiar, R., Papini, P.L.: The lattice theoretic structure of the sets of triangular norms and semi-copulas. Nonlinear Analysis 69, 46–52 (2008)

Reversibility of Triple I Method Based on Łukasiewicz and Goguen Implication

Fu-xi Deng and Jun Li

College of Science, Lanzhou University of Technology, Lanzhou 730050, China
e-mail: dengfuxi@gmail.com, lijun@lut.cn

Abstract. The reversibility of fuzzy inference methods is an important criterion to judge the effect of implication operators matching inference methods. Only the implication operators perfectly match inference methods can fuzzy reasoning have good effect. Furthermore, a fuzzy inference method satisfying reversibility is consistent with the classical boolean logic calculus. In this paper, first the properties of Łukasiewicz and Goguen implication operators are investigated, then necessary and sufficient conditions of reversibility of Triple I Method for FMP and FMT based on them are proved respectively.

Keywords: Triple I Method; Reversibility, Łukasiewicz implication, Goguen implication, FMT (Fuzzy Modus Ponens), FMT (Fuzzy Modus Tollens).

1 Introduction

The methods of fuzzy reasoning play important roles in the design and analysis of fuzzy control and expert systems. FMP (Fuzzy Modus Ponens) and FMT (Fuzzy Modus Tollens) problems are two most important inference models for fuzzy reasoning:

$$FMP: Given\ A \to B\ (rule)\ and\ A^*\ (input), to\ compute\ B^*(output), \quad (1)$$

$$FMT: Given\ A \to B\ (rule)\ and\ B^*\ (input), to\ compute\ A^*(output). \quad (2)$$

Where $A, A^* \in F(X)$ (the set of all fuzzy subsets of universe X) and $B, B^* \in F(Y)$ (the set of all fuzzy subsets of universe Y). Zadeh proposed CRI [2] (Compositional Rule of Inference) method for FMP problem in 1973. After that, many research have been done both theoretically and practically. For improving the CRI method, Wang proposed a new method of fuzzy inference called Triple I (the abbreviation for triple implication) Method in [3,4]. The basic principle is

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 227–232.
springerlink.com

$$(A(x) \to B(y)) \to (A^*(x) \to B^*(y)) \tag{3}$$

should take its maximum whenever $x \in X$ and $y \in Y$.

In [5,6] the unified form of triple I method was formulated based on all left-continuous t-norms and their residuum. In [5,12,13,14,15] reversibility of triple I method were investigated on the account of that a method of fuzzy inference is consistent if it has reversibility (for FMP, reversibility means B^* should be B if $A^* = A$. for FMT, it means A^* should returns to A if $B^* = B$). In these papers, some sufficient conditions of Triple I method for FMP and FMT were obtained based on various triangular norms and fuzzy implication operators.

In the present paper, we investigate the properties of Łukasiewicz t-norm and its residuum, and explore the triple I method based on them. Then we give sufficient and necessary conditions of reversibility of triple I methods for FMP and FMT methods and examine their connection with the conclusions of previous literatures.

2 Preliminaries

Although there are Triple I Method based on other fuzzy implications and Triple I Method whose three implication operators in (3) may be different, the unified form of Triple I Method formulated in [5,6] employs same implication operator in (3). And the implication operators it employs is *regular implication operators*, i.e., the residuum of left-continuous t-norms on which we won't emphasize in the present paper.

PRINCIPLE OF TRIPLE I FOR FMP *(See [3].)The FMP conclusion B^* of (1) is the smallest fuzzy subset of Y which maximize (3).*

Theorem 1. (Triple I Method for FMP) *(See [5].) Suppose the implication operator $\to$ in FMP (1) is regular,then the FMP conclusion of (1) satisfying the PRINCIPLE OF TRIPLE I FOR FMP is*

$$B^*(y) = \sup_{x \in X}\{A^*(x) \otimes (A(x) \to B(y))\}, \quad y \in Y. \tag{4}$$

where $\otimes$ is the t-norm residuated to $\to$.

PRINCIPLE OF TRIPLE I FOR FMT *(See [3].)The FMT conclusion A^* of (2) is the greatest fuzzy subset of X which maximize (3).*

Theorem 2. (Triple I Method for FMT) *(See [6].) Suppose the implication operator $\to$ in FMT (2) is regular,then the FMT conclusion of (2) satisfying the PRINCIPLE OF TRIPLE I FOR FMT can be expressed as follows*

$$A^*(y) = \inf_{y \in Y}\{(A(x) \to B(y)) \to B^*(y)\}, \quad x \in X. \tag{5}$$

The most fundamental deduction rule in logic is modus ponens, it says that if $A \to B$ and A are given, then B follows. Accordingly, it is very natural to require the FMP conclusion B^* in (1) should return to B if the input A^* is A. Similarily, the FMT conclusion A^* in (2) should return to A if the input B^* is B.

There is a sufficient condition of reversibility of Triple I Method for FMP in [5].

Theorem 3. *The Triple I Method for FMP is reversible for normal input, i.e., if the fuzzy subset A^* in (1) and (4) equals to A and A is a normal fuzzy subset of X, then B^* equals to B, where R is a regular implication operator.*

There also is a sufficient condition of reversibility of Triple I Method for FMT in [5].

Theorem 4. *([5]) The Triple I Method for FMT is reversible for co-normal input, i.e., if the fuzzy subset B^* in (2) and (5) equals to B and B' is a normal fuzzy subset of Y, then A^* equals to A, where R is a normal implication operator. A normal implication operator is a regular one who satisfies $a' \to b' = b \to a$ where $a' = 1 - a, b' = 1 - b,\ a, b \in [0, 1]$.*

3 Reversibility of Triple I Method

Firstly, we give the Łukasiewicz Implication operator and t-norm together with their properties we need in this paper. Below,

$$a \otimes_{Ł} b = (a + b - 1) \vee 0, \quad a \to_{Ł} b = (1 - a + b) \wedge 1, \quad \textit{where} \quad a, b, c \in [0, 1].$$

are the Łukasiewicz t-norm and implication operator. They satisfy the following properties:

$$a \otimes_{Ł} (a \to_{Ł} b) = a \wedge b, \quad (a \to_{Ł} b) \to_{Ł} b = a \vee b,\ a, b \in [0, 1]. \tag{6}$$

Theorem 5. *The Triple I Method for FMP based on Łukasiewicz Implication is reversible if and only if*

$$\sup_{x \in X}\{A(x)\} \geq \sup_{y \in Y}\{B(y)\}. \tag{7}$$

Proof. Suppose the triple I method for FMP based on Łukasiewicz Implication is reversible, then the fuzzy subset A^* in (4) equals A when B^* equals B, i.e.,

$$B^*(y) = \sup_{x \in X}\{A(x) \otimes_{Ł} (A(x) \to_{Ł} B(y))\} = B(y), \quad y \in Y. \tag{8}$$

It follows the properties of (6) and (8) that

$$\begin{aligned}&\sup_{x\in X}\{A(x)\otimes_{Ł}(A(x)\to_{Ł}B(y))\}\\&=\sup_{x\in X}\{A(x)\wedge B(y)\}\\&=\sup_{x\in X}\{A(x)\}\wedge B(y)\\&=B(y),\quad y\in Y.\end{aligned}$$

Therefore

$$\sup_{x\in X}\{A(x)\}\geq\sup_{y\in Y}\{B(y)\}.$$

Conversely, suppose (7) holds, then

$$\sup_{x\in X}\{A(x)\}\geq B(y),\quad y\in Y. \tag{9}$$

Again, by the properties of (6), it follows from the Triple I Method for FMP and (9) that

$$\begin{aligned}B^*(y)&=\sup_{x\in X}\{A^*(x)\otimes_{Ł}(A(x)\to_{Ł}B(y))\}\\&=\sup_{x\in X}\{A(x)\otimes_{Ł}(A(x)\to_{Ł}B(y))\}\\&=\sup_{x\in X}\{A(x)\wedge B(y)\}\\&=\sup_{x\in X}\{A(x)\}\wedge B(y)\\&=B(y)\end{aligned}$$

holds for any $y\in Y$ if $A^*=A$ (i.e., B^* returns to B if $A^*=A$), i.e., reversible.

This theorem gives a sufficient and necessary condition of Triple I Method for FMP based on Łukasiewicz Implication. In the following theorem we give a sufficient and necessary condition of Triple I Method for FMT based on the same operator.

Theorem 6. *The Triple I Method for FMT based on Łukasiewicz Implication is reversible if and only if*

$$\inf_{x\in X}\{A(x)\}\geq\inf_{y\in Y}\{B(y)\}. \tag{10}$$

Proof. Suppose the fuzzy subset B^* equals to B when A^* equals A (reversible), i.e.,

$$A^*(y)=\inf_{y\in Y}\{(A(x)\to_{Ł}B(y))\to_{Ł}B(y)\}=A(x),\quad x\in X. \tag{11}$$

It follows property in (6) that

$$\inf_{y\in Y}\{(A(x)\to_{Ł} B(y))\to_{Ł} B(y)\} = \inf_{y\in Y}\{A(x)\vee B(y)\}, \quad x\in X. \tag{12}$$

From (11) and (12) we get

$$A(x) = \inf_{y\in Y}\{A(x)\vee B(y)\} = \inf_{y\in Y}\{A(x)\}\vee B(y), \quad x\in X.$$

Therefore

$$\inf_{x\in X}\{A(x)\} \geq \inf_{y\in Y}\{B(y)\}.$$

Conversely, suppose (10) holds, then for any $x \in X, A(x) \geq \inf_{y\in Y}\{B(y)\}$. When $B^* = B$, by the Triple I Method for FMT and property of $\to_{Ł}$, we have

$$\begin{aligned} A^*(x) &= \inf_{y\in Y}\{(A(x)\to_{Ł} B^*(y))\to_{Ł} B^*(y)\} \\ &= \inf_{y\in Y}\{(A(x)\to_{Ł} B(y))\to_{Ł} B(y)\} \\ &= \inf_{y\in Y}\{A(x)\vee B(y)\} \\ &= A(x)\vee \inf_{y\in Y}\{B(y)\} \\ &= A(x),\ x\in X, \end{aligned}$$

i.e., A^* returns to A. The Triple I Method for FMT based on Łukasiewicz implication is reversible.

Remark 1. The equation in (6), theorem 5 and 6 will still hold if the Łukasiewicz implication is replaced by the Goguen implication (the related Łukasiewicz t-norm will be replaced by product t-norm correspondingly). The proof is analogous and we will not repeat it.

4 Conclusion

The conclusion about reversibility in [5] are based on all regular implication operator, i.e., residuum of left-continuous t-norms. In [12, 13, 14, 15], sufficient conditions of reversibility of Triple I Method were proved based on some specified fuzzy implication. We give two sufficient and necessary conditions of Triple I Method in this paper. This is a further, and at the same time restricted to Łukasiewicz and Goguen implications, development of previous literatures. We will continue our investigation in reversibility of Triple I Method based on other implication operators in later research.

Acknowledgements. This paper was supported by the National Science Foundation of China (No:10771129), the Outstanding Youth Foundation of Lanzhou University of Technology and the Doctoral Foudation of Lanzhou University of Technology.

References

1. Hájek, P.: Metamathematics of fyzzy logic. Kluwer Academic Publishers, Dordrecht (1998)
2. Zadeh, L.A.: Outline of new approach to the analysis of complex systems and decision processes. IEEE Trans Systems Man Cybernetic 3(1), 28–44 (1973)
3. Wang, G.J.: The full implication triple I method for fuzzy reasoning. Science in China (Ser.E) 29, 43–53 (1999) (in Chinese)
4. Wang, G.J.: On the logic foundation of fuzzy reasoning. Information Sciences 117(1-2), 47–88 (1999)
5. Wang, G.J., Fu, L.: Unified forms of triple I methods. Comput. Math. Appl. 49, 923–932 (2005)
6. Liu, H.W., Wang, G.J.: A note on the unified forms of triple I method. Comput. Math. Appl. 52(10-11), 1609–1613 (2006)
7. Liu, H.W., Wang, G.J.: Triple I method based on pointwise sustaining degrees. Comput. Math. Appl. 55, 2680–2688 (2008)
8. Li, H.X.: Probabilistic expression of fuzzy system. Science in China (Ser.F) 36, 373–397 (2006) (in Chinese)
9. Li, J., Wang, G.J.: Formalized solutions for problems of generalized Mudus Ponens under theory of sustentation degree. Acta Electronica Sinica 36, 2190–2194 (2008)
10. Pei, D.W.: On optimal solution of GMP and GMT. Fuzzy Systems and Mathematics 17, 1–9 (2003) (in Chinese)
11. Song, S.J., Feng, C.B., Lee, E.S.: Triple I Method of fuzzy reasoning. Comput. Math. Appl. 44(12), 1567–1579 (2002)
12. Pei, D.W.: Two Triple I Methods for FMT problem and their Reductivity. Fuzzy Systems and Mathematics 15(4), 1–7 (2001) (in Chinese)
13. Pei, D.W.: The full implication triple I method and its reversibility for fuzzy reasoning. Journal of Mathematical Research and Exposition 2, 359–368 (2004) (in Chinese)
14. Hou, J., Li, H.X.: Reductivity of some fuzzy inference methods. Fuzzy Systems and Mathematics 19(4), 90–95 (2005) (in Chinese)
15. Peng, J.Y.: Fuzzy entropy Triple I Methods for FMP and FMT problems and their Reductivity. Systems Engineering Theory & Practice 4, 76–82 (2005) (in Chinese)

Random Truth Theory of Proposition Logic and Its Application

Jia-lu Zhang

Department of Mathematics, Xiangnan University, Chenzhou, Hunan, 423000, China
e-mail: zjl0735@163.com

Abstract. If there is a probability distribution on the valuation domain [0, 1] of logic formulas, the concept of random valuation is introduced, and it is showed that every logic formula determines correspondingly a random function over a probability space. The concept of truth degree of a logic formula, similarity degree and pseudo-metric among two logic formulas are introduced, and it is proved that the truth degree set of all logic formulas and the random logic pseudo-metric space have not isolated point. Based on random truth degree theory three diverse approximate reasoning ways are proposed.

Keywords: Probability space; logic formula; random valuation; random truth degree; random logic pseudo-metric; approximate reasoning.

1 Introduction

Artificial Intelligence, which put emphasis on formalized logic deduction, plays important roles in the subject areas such as logic programming, automatic theorem proving and knowledge reasoning, etc. While numerical computing seems to be completely different and far from the formal deduction method. Hence how to combine these two opposite methods is an attractive research problem. In dealing with combination of logic deduction with probabilistic computing, graded method of logic deduction is extensively used. Because it realized to be the exact means featuring human thinking of which logic deduction and numerical estimation are naturally mixed [1-7]. For example, ref. [6] proposed a thoroughly graded theory of logic deduction where numerical calculations were used throughout the paper, and ref. [7] proposed a graded method of logic deduction based on the concept of similarities where two nice complete theorems were proved in purely formalized ways.

Quantitative logic, proposed by Wang, link up the artificial intelligence and the theories of numerical computing by grading the logic concepts [8-13]. The

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 233–244.
springerlink.com

graded idea and method in quantitative logic followed and developed the corresponding method in ref. [5-7], and proposed different theories of approximate reasoning both in two valued logic, many valued logic and continuous valued logic of which the concepts of truth degree of formulas and pseudo-metric among formulas was constructed. In grading logic concepts in quantitative logic they base on a hidden fact that an atomic formula takes every value in valued domain is equal possible, from the view of probability, there is an equal probability distribution on valued domain and then every atomic formula q determines correspondingly an evenly distribution random variable. This equal possibility of atomic formula valued in valuation space seems to be in conflict with the randomness of atomic formula valued, and hence it also is the flaw of quantitative logic that take no account of randomness.

The rest of this paper is organized as follows. Section 2 illustrates that every logic formula determines correspondingly a random function. In Section 3, the concept of random truth degree is introduced. Section 4 establishes a logic pseudo-metric space. Section 5 presents three diverse approximate reasoning ways. Section 6 concludes the study.

2 Random Function Derived from Logic Formulas

Suppose that ξ is a random variable from a probability space $(\Theta,\, P_0)$ to the unit interval $I=[0,\,1]$, then its distribution function $F_\xi(x)$ has the following form: when $x\le 0$, $F_\xi(x)=0$; and $x\ge 1$, $F_\xi(x)=1$. Hence the distribution function $F_\xi(x)$ determines another probability measure space $(I,\,\mathcal{B},\,P_\xi)$: $\forall B\in\mathcal{B},\ P_\xi(B)=P_0(\{\omega\,|\,\xi(\omega)\in B\})$, where $\mathcal{B}$ be the Borel set family in the unit interval $[0,\,1]$ (see ref. [14]).

Definition 1. *([14]) Let* $(I_n,\,\mathcal{B}_n,\,P_n)=(I,\,\mathcal{B},\,P_\xi)$, $n=1,\,2,\,\cdots$, $I^\infty=\prod_{n=1}^{\infty} I_n$*. Then* $\prod_{n=1}^{\infty}\mathcal{B}_n$ *generates a* $\sigma-$*algebra* $\mathcal{A}$ *on* I^∞*, and there exists a unique probability measure on* $\mathcal{A}$ *such that for any measurable subset* E *of* $\prod_{n=1}^{m} I_n$,

$$P(E\times\prod_{n=m+1}^{\infty} I_n)=(P_1\times\cdots\times P_m)(E),\ m=1,\,2,\,\cdots.$$

P is called the product measure on I^∞. In the following, we also denote I^∞ by Ω, and the probability measurable space $(\Omega,\,\mathcal{A},\,P)$ will often be abbreviated to Ω.

Suppose that $\neg,\,\vee,\,\rightarrow$ are independent logic connectives, $S=\{q_1,\,q_2,\,\cdots\}$ and $F(S)$ is the free algebra of type $(\neg,\,\vee,\,\rightarrow)$ generated by S, i.e., $F(S)$ is the set consisting of all logic formulas, define on I three operations as follows:

$$\neg x = 1 - x,\ x \vee y = max\{x,\ y\},\ x \rightarrow y = R(x,\ y),\ x, y \in I\,,$$

where $R : I^2 \rightarrow I$ is an implication operator. Then I becomes an algebra of type $(\neg,\ \vee,\ \rightarrow)$. If there is a probability distribution on I, i.e., there is a random variable from a probability space to I, then I is called a random unit interval. A random valuation v of $F(S)$ into a random unit interval I is a homomorphism $\nu : F(S) \rightarrow I$ of type $(\neg,\ \vee,\ \rightarrow)$, $\nu(A)$ is the random valuation of A w.r.t. ν. The set consisting of all random valuations of $F(S)$ is denoted by $\Omega(R)$.

A random valuation $\nu : F(S) \rightarrow I$ is uniquely decided by its restriction on S; in other words, every mapping $\nu_0 : S \rightarrow I$ can uniquely be extended to be a random valuation because $F(S)$ is a free algebra generated by S, i.e., if $\nu(q_k) = \nu_k(k = 1,\ 2,\ \cdots)$, then $v = (\nu_1,\ \nu_2,\ \cdots) \in \Omega$. Conversely, if $v = (\nu_1,\ \nu_2,\ \cdots) \in \Omega$, then there exists a unique $\nu \in \Omega(R)$ such that $\nu(q_k) = \nu_k(k = 1,\ 2,\ \cdots)$. This shows that $\varphi : \Omega(R) \rightarrow \Omega$, $\varphi(\nu) = v$ is a bijection. Hence the probability P on Ω can be transferred into $\Omega(R)$ by means of φ, i.e., $P^*(\Sigma) = P(\varphi(\Sigma))$ for any $\Sigma \subseteq \Omega(R)$, where $\varphi(\Sigma)$ is desired a P-measurable set, i.e., $\varphi(\Sigma) \in \mathcal{A}$, φ is called the measured mapping of $\Omega(R)$ and its inverse mapping is denoted by φ^{-1}. We employ $\mathcal{A}^*$ denote the set family of inverse image of all sets in $\mathcal{A}$ under mapping φ, i.e., $\mathcal{A}^* = \{\Sigma \mid \Sigma \subseteq \Omega(R),\ \varphi(\Sigma) \in \mathcal{A}\}$. Thus $(\Omega(R),\ \mathcal{A}^*,\ P^*)$ is also a probability measure space, and the computing with respect to P^* can be transferred into the computing with respect to P via equation $P^*(\Sigma) = P(\varphi(\Sigma))$.

Let $A \in F(S)$ and ξ be a random variable from a probability space $(\Theta,\ P_0)$ to the unit interval I. Then A uniquely defines a random function $A_\xi : \Omega \rightarrow I$ as follows:

$$A_\xi(v) = \varphi^{-1}(v)(A) = \nu(A),\ v \in \Omega\,.$$

In particular, every atomic formula q_i $(i = 1,\ 2,\ \cdots)$ determines a random variable over $(\Omega,\ \mathcal{A},\ P)$, which is denoted by ξ_i $(i = 1,\ 2,\ \cdots)$. These random variables are mutually independent and has the same distribution with ξ. If $A = A(q_{i_1},\ \cdots,\ q_{i_t})$ is a formula consisting of t atomic formulas, then because of $v(q_{i_j})(1 \leq j \leq t)$ taking any value in I, $A_\xi(\cdot)$ is really a t variable function $A_\xi(\xi_{i_1},\ \cdots,\ \xi_{i_t})$ on I^t, and the way $A_\xi(\xi_{i_1},\ \cdots,\ \xi_{i_t})$ acts on $\xi_{i_1},\ \cdots,\ \xi_{i_t}$ in I through $\neg,\ \vee,\ \rightarrow$ is the same as the way acts on $q_{i_1},\ \cdots,\ q_{i_t}$ in $F(S)$ through $\neg,\ \vee,\ \rightarrow$. For example, if $A = \neg q_1 \vee q_2 \rightarrow q_3$, then $A_\xi(\xi_1,\ \xi_2,\ \xi_3) = R(1 - \xi_1,\ \xi_2,\ \xi_3)$.

3 Random Truth Degree of Logic Formulas

Definition 2. *Suppose that ξ is a random variable from a probability space $(\Theta,\ P_0)$ to the unit interval I and its distribution function is $F_\xi(x)$. Let $A = A(q_{i_1},\ \cdots,\ q_{i_t}) \in F(S)$ and $\triangle_t = [0,1]^t$. Then*

$$\tau_\xi^{(R)}(A) = \int_{\Omega(R)} A_\xi(\nu)dP^* = \int_\Omega A_\xi(v)dP$$

$$= \int_{\triangle_t} A_\xi(x_{i_1}, \cdots, x_{i_t}) dF_\xi(x_{i_1}) \cdots dF_\xi(x_{i_t})$$

is called the random truth of formula A w.r.t. the random variable ξ and the implication operator R. $\tau_\xi^{(R)}(A)$ may be abbreviated to $\tau(A)$ if no confusion arises.

Remark 1. (1) The random truth $\tau_\xi^{(R)}(A)$ of A is really the mathematical expectation of random function A_ξ.

(2) If ξ is an evenly distributed random variable on I, then the random truth degree $\tau_\xi^{(R)}(A)$ of A degenerate into the integral truth degree of A in ref. [3,9].

(3) If ξ is an equal probability distributed random variable on $\{0, \frac{1}{n-1}, \cdots, \frac{n-2}{n-1}, 1\}$ and the implication operator is R_{Lu}, then the random truth degree $\tau_\xi^{(R)}(A)$ of A degenerate into the n-valued truth degree of A in ref.[10], and when $n = 2$ it exactly is the truth degree of two valued logic in ref. [11,12].

(4) If ξ has the following probability distribution on $\{0, \frac{1}{n-1}, \cdots, \frac{n-2}{n-1}, 1\}$: $P(\{\xi = 0\}) = \frac{n-1}{n}$, $P(\{\xi = \frac{1}{n-1}\}) = \cdots = P(\{\xi = \frac{n-2}{n-1}\}) = 0$, $P(\{\xi = 1\}) = \frac{1}{n-1}$, and the implication operator is R_{Lu}, then the random truth degree $\tau_\xi^{(R)}(A)$ of A degenerate into the n-valued truth degree of A in ref. [13].

In the following we only consider the implication operator R_{Lu}.

Example 1. Let $A = q_1$, $B = q_1 \to q_2$, $C = q_1 \wedge q_2 \wedge q_3$.

(1) Suppose that ξ has two-point distribution: $P(\{\xi = 1\}) = 0.4$, $P(\{\xi = 0\}) = 0.6$. Computer the truth degree of A, B, C.

(2) Suppose that η has the following probability density function:

$$f(x) = \begin{cases} 2x\,, & 0 \le x \le 1, \\ 0\,, & \text{otherwise}\,. \end{cases}$$

Computer the truth degree of A, B, C.

Solution 1. (1) The distribution of A_ξ as follows: $P(\{A_\xi = 1\}) = P(\{\xi = 1\}) = 0.4$, $P(A_\xi = 0) = 0.6$. Hence $\tau_\xi(A) = \sum_{i=0}^{1} i \times P(\{A_\xi = i\}) = 0.4\,.$

The distribution of B_ξ as follows: $P(\{B_\xi(\xi_1, \xi_2) = 1\}) = P(\{\xi_1 \to \xi_2 = 1\}) = P(\{\xi_1 = 0\})P(\{\xi_2 = 0\}) + P(\{\xi_1 = 0\})P(\{\xi_2 = 1\}) + P(\{\xi_1 = 1\})P(\{\xi_2 = 1\}) = 0.6 \times 0.6 + 0.6 \times 0.4 + 0.4 \times 0.4 = 0.76$, $P(\{B_\xi(\xi_1, \xi_2) = 0\}) = P(\{\xi_1 \to \xi_2 = 0\}) = P(\{\xi_1 = 1\})P(\{\xi_2 = 0\}) = 0.4 \times 0.6 = 0.24$. Hence $\tau_\xi(B) = \sum_{i=0}^{1} i \times P(\{B_\xi(\xi_1, \xi_2) = i\}) = 0.76\,.$

The distribution of C_ξ as follows: $P(\{C_\xi(\xi_1, \xi_2, \xi_3) = 1\}) = P(\{\xi_1 \wedge \xi_2 \wedge \xi_3 = 1\}) = P(\{\xi_1 = 1\})P(\{\xi_2 = 1\})P(\{\xi_3 = 1\}) = 0.4 \times 0.4 \times 0.4 = 0.064$,

$P(\{B_\xi(\xi_1, \xi_2, \xi_3) = 0\}) = 1 - P(\{B_\xi(\xi_1, \xi_2, \xi_3) = 1\}) = 1 - 0.064 = 0.936.$

Hence $\tau_\xi(C) = \sum_{i=0}^{1} i \times P(\{C_\xi(\xi_1, \xi_2, \xi_3) = i\}) = 0.064\,.$

(2) The distribution function of η is

$$F_\eta(x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} x^2\,, & 0 \le x \le 1, \\ 0\,, & x < 0, \\ 1\,, & x > 1, \end{cases}$$

$$\tau_\eta(A) = \int_{\triangle_1} A_\eta(x_1)dF_\eta(x_1) = \int_0^1 x_1 dF_\eta(x_1) = \int_0^1 x_1 dx_1 = \frac{2}{3}\,.$$

Denote $\triangle_2^{(1)} = \{(x_1, x_2)\,|\,0 \le x_1, x_2 \le 1, x_1 \ge x_2\}$, $\triangle_2^{(2)} = \{(x_1, x_2)\,|\,0 \le x_1, x_2 \le 1, x_1 < x_2\}$. Then

$$\tau_\eta(B) = \int_{\triangle_2} B_\eta(x_1, x_2)dF_\eta(x_1)dF_\eta(x_2) = \int_{\triangle_2} ((1-x_1+x_2)\wedge 1)f(x_1)f(x_2)dx_1dx_2$$

$$= \int_{\triangle_2^{(2)}} 2x_1 \cdot 2x_2 dx_1 dx_2 + \int_{\triangle_2^{(1)}} (1-x_1+x_2)\cdot 2x_1 \cdot 2x_2 dx_1 dx_2$$

$$= 4\int_0^1 dx_2 \int_0^{x_2} x_1 x_2 dx_1 + 4\int_0^1 dx_1 \int_0^{x_2} (1-x_1+x_2)x_1 x_2 dx_2 = \frac{1}{2} + \frac{11}{30} = \frac{13}{15}\,.$$

Denote $\triangle_3^{(1)} = \{(x_1, x_2, x_3)\,|\,0 \le x_1, x_2, x_3 \le 1, x_1 \ge x_2 \ge x_3\}$. Then

$$\tau_\eta(C) = \int_{\triangle_2} C_\eta(x_1, x_2, x_3)dF_\eta(x_1)dF_\eta(x_2)dF_\eta(x_3)$$

$$= \int_{\triangle_3} (x_1 \wedge x_2 \wedge x_3)f(x_1)f(x_2)f(x_3)dx_1dx_2dx_3$$

$$= 8\int_{\triangle_3} (x_1 \wedge x_2 \wedge x_3)x_1x_2x_3dx_1dx_2dx_3 = 8 \times 4 \int_{\triangle_3^{(1)}} x_1x_2x_3^2 dx_1dx_2dx_3$$

$$= 32\int_0^1 dx_1 \int_0^{x_1} dx_2 \int_0^{x_2} x_1x_2x_3^2 dx_3 = \frac{32}{105}\,.$$

Proposition 1. *If $A \in F(S)$ is a tautology, then $\tau(A) = 1$.*

Proof. Suppose that $A = A(q_{i_1}, \cdots, q_{i_t})$ is a tautology. Then $A_\xi(v) = 1$ for all $v \in \Omega$, and hence $\tau(A) = \int_\Omega A_\xi(v)dP = 1$.

Proposition 2. *Let $A \in F(S)$. Then $\tau(\neg A) = 1 - \tau(A)$.*

Proposition 3. *Let $A, B \in F(S)$. Then $\tau(A \vee B) = \tau(A) + \tau(B) - \tau(A \wedge B)$.*

Proof. Since $A_\xi \vee B_\xi = A_\xi + B_\xi - A_\xi \wedge B_\xi$. we know that the conclusion holds.

Lemma 1. *Suppose that distribution function $F_\xi(x)$ of ξ is a same order or lower order infinitesimal with x^α for some $\alpha > 0$ when $x \to 0$. Then $\forall \varepsilon > 0$, there is $A \in F(S)$ such that $0 < \tau(A) < \varepsilon$.*

Proof. Since the value domain of ξ is [0,1], $\lim\limits_{x_i \to 0} F_\xi(x_i) = \lim\limits_{x_i \to -\infty} F_\xi(x_i) = 0$. Let $A_t = q_1 \wedge \cdots \wedge q_t$, where $q_i \in S$ and $q_i \neq q_j (i,j = 1, \cdots, t)$ when $i \neq j$. Obviously $\tau(A) > 0$. In the following we prove that

$$\lim_{t \to \infty} \tau(A_t) = \lim_{t \to \infty} \tau(q_1 \wedge \cdots \wedge q_t) = 0\,,$$

then $\tau(A) < \varepsilon$. In fact, denote

$$\delta_t = \{(x_1, \cdots, x_t) \in \triangle_t \,|\, x_i > (\frac{1}{t})^{\frac{\alpha}{2}},\, i = 1, \cdots, t\}\,,$$

then $x_1 \wedge \cdots \wedge x_t \leq (\frac{1}{t})^{\frac{\alpha}{2}}$ on $\triangle_t - \delta_t$. Thus

$$\tau(q_1 \wedge \cdots \wedge q_t) = \int_{\triangle_t} x_1 \wedge \cdots \wedge x_t dF_\xi(\omega_t)$$

$$= \int_{\triangle_t - \delta_t} x_1 \wedge \cdots \wedge x_t dF_\xi(\omega_t) + \int_{\delta_t} x_1 \wedge \cdots \wedge x_t dF_\xi(\omega_t)$$

$$\leq (\frac{1}{t})^{\frac{\alpha}{2}} + (1 - F_\xi((\frac{1}{t})^{\frac{\alpha}{2}}))^t\,.$$

Because $\lim\limits_{t \to \infty} (\frac{1}{t})^{\frac{\alpha}{2}} = 0$ and $F_\xi(x)$ is a same order or lower order infinitesimal with $x^\alpha (\alpha > 0)$ when $x \to 0$, $\lim\limits_{t \to \infty} \frac{F_\xi((\frac{1}{t})^{\frac{\alpha}{2}})}{(\frac{1}{t})^{\frac{\alpha}{2}}}$ equals to finite valued or ∞. It follows from

$$(1 - F_\xi((\frac{1}{t})^{\frac{\alpha}{2}}))^t = \{[1 - F_\xi((\frac{1}{t})^{\frac{\alpha}{2}})]^{\frac{1}{F((\frac{1}{t})^{\frac{\alpha}{2}})}}\}^{\frac{F_\xi((\frac{1}{t})^{\frac{\alpha}{2}})}{(\frac{1}{t})^{\frac{\alpha}{2}}} \cdot t^{\frac{\alpha}{2}}}$$

and $\lim\limits_{t \to \infty} (1 - \frac{1}{t})^t = \frac{1}{e}$ that $\lim\limits_{t \to \infty} (1 - F_\xi((\frac{1}{t})^{\frac{\alpha}{2}}))^t = 0$. Therefore we can conclude the conclusion on basis of the above integration inequality.

Lemma 2. *If there is $0 < r < 1$ such that $F_\xi(x) > r$ when x sufficiently small, then for $\varepsilon > 0$ there is $A \in F(S)$ such that $0 < \tau(A) < \varepsilon$.*

Proof. Taking $t_0 \in N$ and $\frac{1}{t_0} < \frac{\varepsilon}{2}$. Let $A_t = q_1 \wedge \cdots \wedge q_t$, where $q_i \in S$ and $q_i \neq q_j (i,j = 1, \cdots, t)$ when $i \neq j$. Obviously $\tau(A_t) > 0$. Denote

$$\delta_{t_0} = \{(x_1, \cdots, x_t) \in \triangle_t \,|\, x_i > \frac{1}{t_0},\, i = 1, \cdots, t\}\,,$$

then $x_1 \wedge \cdots \wedge x_t \leq \dfrac{1}{t_0}$ on $\triangle_t - \delta_{t_0}$. Hence

$$\tau(q_1 \wedge \cdots \wedge q_t) = \int_{\triangle_t} x_1 \wedge \cdots \wedge x_t dF(\omega_t)$$

$$= \int_{\triangle_t - \delta_{t_0}} x_1 \wedge \cdots \wedge x_t dF_\xi(\omega_t) + \int_{\delta_{t_0}} x_1 \wedge \cdots \wedge x_t dF_\xi(\omega_t) \leq \frac{1}{t_0} + \int_{\delta_{t_0}} dF_\xi(\omega_t)$$

$$= \frac{1}{t_0} + (1 - F_\xi(\frac{1}{t_0})) \cdot (1 - F_\xi(\frac{1}{t_0})) \cdot \cdots \cdot (1 - F_\xi(\frac{1}{t_0})) \leq \frac{1}{t_0} + (1 - r)^t\,.$$

Since $0 < r < 1$, there is t_1 such that $(1 - r)^t < \frac{\varepsilon}{2}$ as $t > t_1$. Taking $t = max\{t_0, t_1\}$, then $A = A_t$ is a desired formulas.

Theorem 1. *Suppose that distribution function $F(x)$ of ξ satisfies the condition of Lemma 1 or 2 and $A = A(q_{i_1}, \cdots, q_{i_t}) \in F(S)$. Then for any $\varepsilon > 0$, there is a formulas $B \in F(S)$ such that $0 < |\tau(A) - \tau(B)| < \varepsilon$.*

Proof. Taking $m > max\{i_1, \cdots, i_t\}$, it follows from Lemma 1 or 2 that there is $C = C(l) = q_{m+1} \wedge \cdots \wedge q_{m+l}$ such that $0 < \tau(C) < \varepsilon$. If $\tau(A) = 0$, then taking $B = C(l)$ and we have $|\tau(A) - \tau(B)| < \varepsilon$. If $\tau(A) = 1$, then taking $B = \neg C$ and we have $\tau(B) = 1 - \tau(C) \neq \tau(A)$ and $|\tau(A) - \tau(B)| < \varepsilon$. Hence it is no hurt to assume that $0 < \tau(A) < 1$. In this case, if $\tau(A) > 1 - \varepsilon$, then taking a tautology B and we have $0 < |\tau(A) - \tau(B)| < \varepsilon$. If $0 < \tau(A) \leq 1 - \varepsilon$, then taking $B = \neg A \rightarrow C(l)$ and we have

$$B_\xi = R_{Lu}(1 - A_\xi,\, C_\xi) = (A_\xi + C_\xi) \wedge 1 \geq A_\xi\,.$$

Thus

$$\tau(B) = \int_\Omega B_\xi dP \leq \int_\Omega (A_\xi + C_\xi) dP = \int_\Omega A_\xi dP + \int_\Omega C_\xi dP = \tau(A) + \tau(C)\,.$$

Therefore

$$|\tau(A) - \tau(B)| = \tau(B) - \tau(A) \leq \tau(C) < \varepsilon\,.$$

In the following we prove that $\tau(A) \neq \tau(B)$.

In fact, since $0 < \tau(A) \leq 1 - \varepsilon$ we know that $P\{v :\, A_\xi < 1\} \geq \varepsilon$. From $0 < \tau(C) < \varepsilon$ we also know that $P\{v : C_\xi > 0\} > 0$. Then it follows from that $A_\xi(\xi_{i_1}, \cdots, \xi_{i_t})$ and $C_\xi(\xi_{m+1}, \cdots, \xi_{m+l})$ are independent that

$$P\{v :\, A_\xi < 1,\, C_\xi > 0\} = P\{v :\, A_\xi < 1\} \cdot P\{v : C_\xi > 0\} > 0\,.$$

Because $B_\xi = (A_\xi + C_\xi) \wedge 1 > A_\xi$ on set $\{v :\, A_\xi < 1,\, C_\xi > 0\}$, we have $\tau(A) < \tau(B)$.

4 Similarity Degree and Pseudo-metric among Formulas

Definition 3. *Suppose that ξ is a random variable from a probability space $(\Theta,\, P_0)$ to the unit interval I and $A,\, B \in F(S)$. Define*

$$\mu_{\xi}^{(R)}(A,\, B) = \tau_{\xi}^{(R)}((A \to B) \wedge (B \to A))\,.$$

Then $\mu_{\xi}^{(R)}(A,\, B)$ is called the random similarity degree of A and B based on the random variable ξ and the implication operator R. $\mu_{\xi}^{(R)}(A,\, B)$ may be abbreviated to $\mu(A,\, B)$ if no confusion arises.

Theorem 2. *Let $A,\, B,\, C \in F(S)$. Then*
(1) If $A \approx B$ then $\mu(A,\, B) = 1$.
(2) $\mu(A,\, B) + \mu(B,\, C) \le 1 + \mu(A,\, C)$.

Proof. (1) is obviously.
(2) Denote $f(x,\, y) = R(x,\, y) \wedge R(y,\, x)$, then $f(a,\, c) \ge f(a,\, b) + f(b,\, c) - 1$. Hence

$$\mu(A,\, C) = \int_{\Omega} f(A_\xi,\, C_\xi)dP \ge \int_{\Omega} [f(A_\xi,\, B_\xi) + f(B_\xi,\, C_\xi) - 1]dP$$

$$= \int_{\Omega} f(A_\xi,\, B_\xi)dP + \int_{\Omega} f(B_\xi,\, C_\xi)dP - 1 = \mu(A,\, B) + \mu(B,\, C) - 1\,.$$

Theorem 3. *Suppose that ξ is a random variable from a probability space $(\Theta,\, P_0)$ to the unit interval I and $A,\, B \in F(S)$. Define*

$$\rho_{\xi}^{(R)}(A,\, B) = 1 - \mu_{\xi}^{(R)}(A,\, B)\,.$$

Then $\rho_{\xi}^{(R)}$ is the pseudo-metric on $F(S)$. $(F(S), \rho_{\xi}^{(R)})$ is called the random pseudo-metric based on random variable ξ and implication operator R. $\rho_{\xi}^{(R)}(A,\, B)$ may be abbreviated to $\rho(A,\, B)$ if no confusion arises.

Proof. It follows from Definition 3 and Theorem 2 that $\rho_{\xi}^{(R)}$ is the pseudo-metric on $F(S)$.

Theorem 4. *Suppose that ξ is a random variable from a probability space $(\Theta,\, P_0)$ to the unit interval I. If $R = R_{Lu}$, then*

$$\rho(A,\, B) = \int_{\Omega} |A_\xi(v) - B_\xi(v)|dP\,.$$

Proof. The proof of this theorem is checked simple.

Example 2. Let $A = q_1$, $B = \neg q_1 \to q_2$, Taking random variable ξ, η as in Example 1 and implication operator $R = R_{Lu}$. Computing $\rho_\xi(A, B)$, $\rho_\eta(A, B)$.

Solution 2. Note that

$$(q_1 \to (\neg q_1 \to q_2)) \wedge ((\neg q_1 \to q_2) \to q_1) \approx (\neg q_2 \to q_1) \to q_1 \approx \neg q_2 \vee q_1 \,.$$

Similar to Example 1, we obtain by computing

$$\mu_\xi(A, B) = \tau_\eta(\neg q_2 \vee q_1) = 0.76 \,,$$

$$\mu_\eta(A, B) = \tau_\eta(\neg q_2 \vee q_1) = \int_{\triangle_2} ((1 - x_2) \vee x_1))x_1 x_2 dx_1 dx_2 = 0.5 \,,$$

$$\rho_\xi(A, B) = 1 - \mu_\xi(A, B) = 1 - 0.76 = 0.24 \,,$$

$$\rho_\eta(A, B) = 1 - \mu_\eta(A, B) = 0.5 \,.$$

Theorem 5. *Suppose that distribution function $F(x)$ of ξ satisfies the condition of Theorem 1 and $A = A(q_{i_1}, \cdots, q_{i_t}) \in F(S)$. Then for any $\varepsilon > 0$, there is a formula $B \in F(S)$ such that*

$$0 < \rho(A, B) < \varepsilon \,.$$

Proof. Taking $m > max\{i_1, \cdots, i_t\}$, there is $C = C(l) = q_{m+1} \wedge \cdots \wedge q_{m+l}$ such that $0 < \tau(C) < \varepsilon$ by Theorem 1. If $\tau(A) = 1$, then we take $B = \neg C$ and

$$\rho(A, B) = \int_\Omega |1 - (1 - C_\xi)| dP = \int_\omega \xi_{m+1} \wedge \cdots \wedge \xi_{m+l} dP = \tau(C) < \varepsilon \,.$$

If $\tau(A) < 1$, then we take $B = \neg A \to C$ and

$$\rho(A, B) = \int_\omega |A_\xi - R_{Lu}(1 - A_\xi, C_\xi)| dP = \int_\omega |A_\xi - (A_\xi + C_\xi) \wedge 1| dP$$

$$\leq \int_\omega |(A_\xi + C_\xi) - A_\xi| dP = \int_\omega C_\xi dP = \tau(C) < \varepsilon \,.$$

Similar to the proof of $\tau(A) \neq \tau(B)$ in Theorem 1, we can prove that $\rho(A, B) > 0$.

5 Approximation Reasoning Based on Truth Degrees

In this section, we provide three approximate reasoning ways based on the random truth degrees.

Definition 4. *Suppose that ξ is a random variable from a probability space (Θ, P_0) to the unit interval I, $\Gamma \subset F(S)$, $A \in F(S)$.*

(i) If there are finite formulas $\{B_1, \cdots, B_n\} \subseteq \Gamma$ and $N \subseteq \Omega$, $P(N) = 0$ such that $\forall v \in \Omega - N$, $\bigotimes_{i=1}^{n}(B_i)_\xi(v) \to A_\xi(v) = 1$, where $\bigotimes_{i=1}^{n}(B_i)_\xi(v) = (B_1)_\xi(v) \otimes \cdots \otimes (B_n)_\xi(v)$, then we call A is an a.e. conclusion of Γ. In particular, if $\Gamma = \emptyset$, then we call A is an a.e. theorem.

(ii) If $\forall \varepsilon > 0$, $\delta > 0$, there are finite formulas $\{B_1, \cdots, B_n\} \subseteq \Gamma$ such that $P(\{v : \bigotimes_{i=1}^{n}(B_i)_\xi(v) - A_\xi(v) \geq \varepsilon\}) < \delta$, then we call A is a conclusion of Γ in probability. In particular, if $\Gamma = \emptyset$, then we call A is a theorem in probability.

(iii) If $\forall \varepsilon > 0$ there are finite formulas $\{B_1, \cdots, B_n\} \subseteq \Gamma$ such that $\tau(B_1 \& \cdots \& B_n \to A) > 1 - \varepsilon$, then we call A is a conclusion of Γ in truth degree. In particular, if $\Gamma = \emptyset$, then we call A is a theorem in truth degree.

Remark 2. In Lukasiewicz logic system, if $\Gamma = \{B_1, \cdots, B_n\}$ is finite and A is a conclusion of Γ (see ref. [2,4]), then $\forall \nu \in \Omega(R), \nu(B_1) \otimes \cdots \otimes \nu(B_n) \to \nu(A) = 1$ by ref. [4] Section 8.4 Exercise 1,4. Hence $\forall v \in \Omega, \bigotimes_{i=1}^{n}(B_i)_\xi(v) \to A_\xi(v) = 1$. This shows that A is an a.e. conclusion of Γ.

Theorem 6. *If A is an a.e. conclusion of Γ, then A is a conclusion of Γ in probability.*

Proof. Since A is an a.e. conclusion of Γ, there are formulas $\{B_1, \cdots, B_n\} \subseteq \Gamma$ and $N \subseteq \Omega$, $P(N) = 0$ such that $\{v : \bigotimes_{i=1}^{n}(B_i)_\xi(v) \to A_\xi(v) \neq 1\} \subseteq N$. Taking monotone decreasing sequence $\{\varepsilon_k\}$ and $\lim_{k\to\infty} \varepsilon_k = 0$, then $\{v : \bigotimes_{i=1}^{n}(B_i)_\xi(v) - A_\xi(v) \geq \varepsilon_k\}$ is monotone increasing and

$$\lim_{k\to\infty} \{v : \bigotimes_{i=1}^{n}(B_i)_\xi(v) - A_\xi(v) \geq \varepsilon_k\} = \{v : \bigotimes_{i=1}^{n}(B_i)_\xi(v) \to A_\xi(v) \neq 1\}\,.$$

By the continuity of probability we have that

$$\lim_{k\to\infty} P(\{v : \bigotimes_{i=1}^{n}(B_i)_\xi(v) - A_\xi(v) \geq \varepsilon_k\}) = P(\{v : \bigotimes_{i=1}^{n}(B_i)_\xi(v) \to A_\xi(v) \neq 1\}) = 0\,.$$

Hence for any $\delta > 0$ there is $K_1 > 0$ such that

$$P(\{v : \bigotimes_{i=1}^{n}(B_i)_\xi(v) - A_\xi(v) \geq \varepsilon_k\}) < \delta$$

as $k > K_1$. On the other hand, for any $\varepsilon > 0$ there is also $K_2 > 0$ such that $\varepsilon_k < \varepsilon$ as $k > K_2$. Thus if we take $K = max\{K_1, K_2\}$ then

$$P(\{v : \bigotimes_{i=1}^{n}(B_i)_\xi(v) - A_\xi(v) \geq \varepsilon\}) \leq P(\{v : \bigotimes_{i=1}^{n}(B_i)_\xi(v) - A_\xi(v) \geq \varepsilon_k\}) < \delta$$

as $k > K$. This proves Theorem 6.

Theorem 7. *If A is a conclusion of Γ in probability, then A is a conclusion of Γ in truth degree.*

Proof. For any $\varepsilon > 0$, by the assumption that A is a conclusion of Γ in probability, there are finite formulas $\{B_1, \cdots, B_n\} \subseteq \Gamma$ such that

$$P(\{v : \bigotimes_{i=1}^{n}(B_i)_\xi(v) - A_\xi(v) \geq \frac{\varepsilon}{2}\}) < \frac{\varepsilon}{2}.$$

In the following we denote $\Omega_1 = \{v : \bigotimes_{i=1}^{n}(B_i)_\xi(v) - A_\xi(v) \geq \frac{\varepsilon}{2}\}$. Hence

$$\tau(B_1\& \cdots \&B_n \to A) = \int_{\Omega}[(1-\bigotimes_{i=1}^{n}(B_i)_\xi(v)+A_\xi(v))\wedge 1]dP$$

$$= \int_{\Omega_1}(1-\bigotimes_{i=1}^{n}(B_i)_\xi(v)+A_\xi(v))dP + \int_{\Omega-\Omega_1}[(1-\bigotimes_{i=1}^{n}(B_i)_\xi(v)+A_\xi(v))\wedge 1]dP$$

$$\geq \int_{\Omega-\Omega_1}(1-\frac{\varepsilon}{2})dP = (1-\frac{\varepsilon}{2})P(\Omega-\Omega_1) \geq (1-\frac{\varepsilon}{2})^2 = 1-\varepsilon+\frac{\varepsilon^2}{4} \geq 1-\varepsilon\,.$$

This proves that A is a conclusion of Γ in truth degree.

6 Conclusion

In this paper, by investigating the valuation of logic formulas under the circumstances that there is a probability distribution on the valued domain of logic formulas, we extend the concepts of truth degree, similarity degree and pseudo-metric in quantitative logic to random truth degree, random similarity degree and random pseudo-metric respectively. It is proved that the truth degree set of all logic formulas and the random logic pseudo-metric space have not isolated point. Three diverse approximate reasoning ways are proposed based on random truth degree. The study of this paper and the further works may provide a more flexible logic reasoning theory in artificial intelligence.

Acknowledgements. The work of this paper has been supported by the construction program of the key discipline in Hunan Province and the aid program for science and technology innovative research team in higher educational institutions of Hunan Province.

References

1. Esteva, F., Godo, L.: Monoidal t-norm based logic: towards a logic for left-continuous t-norms. Fuzzy Sets and Systems 124(3), 271–288 (2001)
2. Hájek, P.: Metamathematics of fuzzy logic. Kluwer Adacemic Publishers, Dordrecht (1998)
3. Wang, G.J.: Non-classical logic and approximate Reasoning. Science Press, Beijing (2000)
4. Wang, G.J.: Introduction to mathematical logic and resolution principle. Science Press, Beijing (2003)
5. Dubois, D., Prad, H.: Possibility theory, probability theory and multiple-valued logics. Ann. Math. Artif. Intell. 32, 35–66 (2001)
6. Pavalk, J.: On fuzy logic (I, II, III). Z Math. Logik Grundlagen Math. 25, 45–52, 119–134, 447–464 (1979)
7. Ying, M.S.: A logic for approximate reasoning. J. Symb. Log. 59, 830–837 (1994)
8. Wang, G.J.: A universal theory of measure and integral on valuation spaces with respect to diverse implication operators. Science in China (Series E) 43(6), 586–594 (2000)
9. Wang, G.J., Leung, Y.: Integrated semantic and logic metric space. Fuzzy Sets and Systems 136(1), 71–91 (2003)
10. Li, B.J., Wang, G.J.: Theory of truth degrees of formulas in Lukasiewicz n-valued propositional logic and a limit theorem. Science in China (Series F) 48(6), 727–736 (2005)
11. Wang, G.J., Wang, W.: Logic metric space. Acta Mathematica Sinica 44(1), 159–168 (2001)
12. Wang, G.J., Fu, L., Song, J.S.: Theory of truth degrees of propositions in two valued proposition logic. Science in China (Series A) 45(9), 1106–1116 (2002)
13. Li, J., Li, S.P., Xia, Y.F.: Theory of truth degrees of propositions in Lukasiewicz n-valued logic. Acta Mathematica Sinica 47(4), 769–780 (2004)
14. Halmos, P.R.: Measure theory. Springer, New York (1974)
15. Adams, E.W.: A primer of probability logic. CSLI Publications, Stanford (1998)

Soft Set Based Approximate Reasoning: A Quantitative Logic Approach

Feng Feng[1,2], Yong-ming Li[1,3], Chang-xing Li[2], and Bang-he Han[1,4]

[1] College of Mathematics and Information Science, Shaanxi Normal University, Xi'an 710062, China
[2] Department of Applied Mathematics, School of Science, Xi'an Institute of Posts and Telecommunications, Xi'an 710121, China
e-mail: fengnix@hotmail.com
[3] College of Computer Science, Shaanxi Normal University, Xi'an 710062, China
e-mail: liyongm@snnu.edu.cn
[4] School of Science, Xidian University, Xi'an 710071, China
e-mail: bhhan@stu.snnu.edu.cn

Abstract. Soft set theory is a newly emerging mathematical approach to vagueness. However, it seems that there is no existing research devoted to the discussion of applying soft sets to approximate reasoning. This paper aims to initiate an approximate reasoning scheme based on soft set theory. We consider proposition logic in the framework of a given soft set. By taking parameters of the underlying soft set as atomic formulas, the concept of (well-formed) formulas over a soft set is defined in a natural way. The semantic meaning of formulas is then given by taking objects of the underlying soft set as valuation functions. We propose the notion of decision soft sets and define decision rules as implicative type of formulas in decision soft sets. Motivated by basic ideas from quantitative logic, we also introduce several measures and preorders to evaluate the soundness of formulas and decision rules in soft sets. Moreover, an interesting example is presented to illustrate all the new concepts and the basic ideas initiated here.

Keywords: Soft Set, approximate reasoning, truth degree, quantitative logic.

1 Introduction

Complex problems involving various vagueness are pervasive in many areas of modern technology. These practical problems arise in such diverse areas as economics, engineering, environmental science, social science, and medical science among others. While a wide range of mathematical disciplines like probability theory, fuzzy set theory [1], rough set theory [2] and interval mathematics [3] are useful mathematical approaches to dealing with vagueness and

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 245–255.
springerlink.com

uncertainty, each of them has its advantages as well as inherent limitations. One major weakness shared by these theories is possibly the inadequacy of the parametrization tool of the theory [4].

In 1999, Molodtsov [4] initiated *soft set theory* as a new mathematical tool for dealing with uncertainty, which seems to be free from the inherent difficulties affecting existing methods. The theory of soft sets has potential applications in various fields like the smoothness of functions, game theory, operations research, Riemann integration, Perron integration, probability theory, and measurement theory [4, 5]. Although soft set theory and its applications has been very active research topic in recent years, we find there is no existing research devoted to the discussion of applying soft sets to approximate reasoning.

Quantitative logic [6] establishes an interesting connection between mathematical logic and numerical computation. The basic idea of quantitative logic is to provide a graded approach to propositional logic. In this paper, motivated by basic ideas from quantitative logic we consider proposition logic in the framework of a given soft set and propose some measures to evaluate the soundness of formulas over a soft set. This naturally gives rise to a new approach to approximate reasoning, which is mainly based on soft set theory and quantitative logic. Moreover, we give some fresh ideas to soft set based decision making by introducing decision rules in soft sets and focusing on the rule induction of decision soft sets.

2 Preliminaries

2.1 The Degree of the Truth of Formulas (By Wang and Zhou [6])

As a first step to grade some important logic concepts, we need to measure to what extent a given formula is true. This motivates us to consider the degree of the truth of formulas in propositional logic systems.

We assume that the reader is familiar with the rudiments of some commonly used propositional logic systems like classical two-valued logic system L, Łukasiewicz many-valued propositional logic systems $Ł_n$ and Łuk, Gödel and Goguen propositional logic systems G and Π, and the R_0 type many-valued propositional logic systems $\mathscr{L}_n^*$. For more detail the reader is referred to [6].

Let $\mathscr{F}(S)$ be the set of all formulas generated by axiomatic formulas in S, and let W be a valuation domain. We know that $\mathscr{F}(S)$ is a free algebra of type $(\neg, \vee, \rightarrow)$. Suppose that $A = A(p_1, \cdots, p_m)$ is a formula built up from the atomic formulas $p_1, \cdots, p_m$ using the logical connectives $\neg$, $\vee$ and $\rightarrow$. If we substitute x_i for p_i ($i \in [m] = \{1, 2, \cdots, m\}$) and interpret the logical connectives as the corresponding operations on W, we can obtain an m-ary function $\bar{A}(x_1, \cdots, x_m) : W^m \rightarrow W$, called the *truth function* induced by A. Using truth functions the degree of the truth of formulas can be defined as follows.

Definition 1. ([6]) *Let $A = A(p_1, \cdots, p_m)$ be a formula containing m atomic formulas $p_1, \cdots, p_m$ in a n-valued propositional logic system, and let $\bar{A}$ be the truth function induced by A. Define*

$$\tau_n(A) = \frac{1}{n^m} \sum_{i=1}^{n-1} \frac{i}{n-1} \mid \bar{A}^{-1}(\frac{i}{n-1}) \mid,$$

where $\mid \cdot \mid$ denotes the cardinality of a set. $\tau_n(A)$ is called the degree of the truth *of A in the n-valued system.*

Notice that in the case of $n = 2$, we have

$$\tau_2(A) = \frac{\mid \bar{A}^{-1}(1) \mid}{2^m}. \tag{1}$$

In this case, the associated valuation domain is $W = \{0, 1\}$. Each 0-1 vector $(x_1, \cdots, x_m) \in W^m$ naturally induces a valuation $v : \mathscr{F}(S) \to \{0, 1\}$ given by $v(p_i) = x_i$ for $i \in [m]$, and $v(p_k) = 0$ for other atomic formulas p_k. The quantity $\tau_2(A)$ expresses the general possibility for the formula A to be true in classical two-valued logic system L. Hence it is natural and reasonable to call $\tau_2(A)$ the degree of the truth of A. Also it is clear that A is a tautology (contradiction) if and only if $\tau_2(A) = 1$ ($\tau_2(A) = 0$). Hence we can say that the notion of the degree of the truth is the result of grading the notion of tautology.

2.2 Soft Set Theory (By Molodtsov [4])

Let us introduce now the notion of soft sets which is a newly-emerging mathematical approach to vagueness.

Let U be an initial universe of objects and E_U (simply denoted by E) the set of parameters in relation to the objects in U. In this study we restrict our discussion to the case that both U and E are nonempty finite sets. By parameters we usually mean attributes, characteristics, or properties of the objects in U. Let $\mathscr{P}(U)$ denote the power set of U.

Definition 2. ([7]) *A pair $S = (F, A)$ is called a soft set over U, where $A \subseteq E$ and $F : A \to \mathscr{P}(U)$ is a set-valued mapping.*

Roughly speaking we can say that soft sets are crisp sets determined by parameters. In other words, a soft set over U is a parameterized family of subsets of the universe U. The absence of any restrictions on the approximate description in soft set theory makes this theory very convenient and easily applicable in practice. We may use any suitable parametrization—with the help of words and sentences, real numbers, functions, mappings, etc.

Definition 3. ([8]) *A soft set $S = (F, A)$ over U is said to be* full *if $\bigcup_{a \in A} F(a) = U$. A full soft set $S = (F, A)$ over U is called a* covering soft set *if $F(a) \neq \emptyset$, $\forall a \in A$.*

Definition 4. *Let (F, A) and (G, B) be two soft sets over U. Then (G, B) is called a* soft subset *of (F, A), denoted by $(G, B) \subseteq (F, A)$, if $B \subseteq A$ and $G(b) \subseteq F(b)$ for all $b \in B$.*

It is worth noting that rough sets and soft sets are distinct but closely related mathematical approaches to vagueness. First one can observe the strong connection between information systems and soft sets.

Definition 5. ([9]) An information system *(or a* knowledge representation system*) is a pair $\mathscr{I} = (U, A)$ of non-empty finite sets U and A, where U is a set of objects and A is a set of attributes; each attribute $a \in A$ is a function $a : U \to V_a$, where V_a is the set of values (called* domain*) of attribute a.*

Let $S = (F, A)$ be a soft set over U. If U and A are both non-empty finite sets, then S could induce an information system $\mathscr{I} = (U, A)$ in a natural way. In fact, for any attribute $a \in A$, one can define a function $a : U \to V_a = \{0, 1\}$ by

$$a(x) = \begin{cases} 1, & \text{if } x \in F(a), \\ 0, & \text{otherwise.} \end{cases}$$

Therefore every soft set may be considered as an information system. This justifies the tabular representation of soft sets widely used in the literature (e.g., see the soft set tables used in what follows). Conversely, it is worth noting that soft sets can also be applied to represent information systems. Let $\mathscr{I} = (U, A)$ be an information system. Taking $B = \bigcup_{a \in A}\{a\} \times V_a$, as the parameter set, then a soft set (F, B) can be defined by setting

$$F(a, v) = \{x \in U : a(x) = v\},$$

where $a \in A$ and $v \in V_a$. By the above discussion, it is easy to see that once given a soft set $S = (F, A)$ over U, we obtain an information system $\mathscr{I} = (U, A)$ corresponding to the soft set S, whence we shall be able to construct rough approximations and discuss Pawlak's rough sets based on the Pawlak approximation space (U, R) induced by the information system $\mathscr{I}$.

Moreover, it is interesting to find that Pawlak's rough set model may be considered as a special case of Molodtsov's soft sets. To see this, suppose that (U, R) is a Pawlak approximation space and $X \subseteq U$. Let $R(X) = (R_*X, R^*X)$ be the rough set of X with respect to R. Consider two predicates $p_1(x), p_2(x)$, which mean "$[x]_R \subseteq X$" and "$[x]_R \cap X \neq \emptyset$", respectively. The predicates $p_1(x), p_2(x)$ may be treated as elements of a parameter set; that is, $E = \{p_1(x), p_2(x)\}$. Then we can define a set-valued mapping

$$F : E \to \mathscr{P}(U),\ p_i(x) \mapsto F(p_i(x)) = \{x \in U : p_i(x)\},$$

where $i = 1, 2$. It follows that the rough set $R(X)$ may be considered a soft set (F, E) with the following representation

$$(F, E) = \{(p_1(x), R_*X), (p_2(x), R^*X)\}.$$

Here we conclude that in rough set theory vague concepts can be interpreted in two different ways, namely the lower and upper approximations; in soft set theory, however, we may interpret a vague concept in a wide variety of distinct ways according to different parameters.

3 Approximate Reasoning Based on Soft Sets

3.1 Formulas over a Soft Set

Let $S = (F, A)$ be a soft set over U. Sometimes we distinguish in S a partition of A into two classes $C, D \subseteq A$ of parameters, called condition and decision (action) parameters, respectively. The tuple $S = (F, C, D)$ is then called a *decision soft set* over U. The soft sets $S_C = (F, C)$ and $S_D = (F, D)$ are called the *condition soft subset* and the *decision soft subset* of S, respectively.

Definition 6. (Formulas over a soft set) *Let $S = (F, A)$ be a soft set over U. Then each parameter $a \in A$ is called an* atomic formula *over S. A finite combination of atomic formulas connected by the logical connectives is a (*well-formed*) formula (also called* proposition*) over the soft set $S = (F, A)$.*

The set of all formulas over a soft set S is denoted by $\mathscr{F}(S)$. Note that different logical connectives might be chosen for different propositional logic systems. For instance, if we use the logical connectives $\neg, \wedge, \rightarrow$, then $\mathscr{F}(S)$ is a free algebra generated by A as follows:

- $A \subseteq \mathscr{F}(S)$;
- If $\varphi \in \mathscr{F}(S)$, then $\neg\varphi \in \mathscr{F}(S)$;
- If $\varphi, \psi \in \mathscr{F}(S)$, then $\varphi \wedge \psi, \varphi \rightarrow \psi \in \mathscr{F}(S)$;
- Every formula in $\mathscr{F}(S)$ is generated by above rules.

We say that formulas in $\mathscr{F}(S)$ are well-formed since they are built properly up from parameters (i.e. atomic formulas) by using certain logical connectives according to the given rules.

For a decision soft set $S = (F, C, D)$ formulas from $\mathscr{F}(S_C)$ and $\mathscr{F}(S_D)$ are sometimes called condition and decision formulas, respectively.

3.2 Valuation and Semantic Interpretation of Formulas

Recall that in propositional logic a valuation is a homomorphism from the free algebra of all formulas into a specific algebra called the valuation domain. The semantics of a propositional logic system provides an effective approach towards evaluating the soundness of formulas by using valuations.

Let us now consider valuation and semantic interpretation of formulas in the framework soft set theory. Given a soft set $S = (F, A)$ over a universe U, we immediately obtain a set of objects, i.e., $V_S = \bigcup_{a \in A} F(a)$. For the sake of convenience we always assume in what follows that the soft set under our consideration is a full one; hence we have $V_S = U$. Now for any object $u \in U$, it is an interesting observation that object u induces a valuation ν_u for formulas over the soft set $S = (F, A)$. Specifically, $\nu_u : \mathscr{F}(S) \to \{0, 1\}$ is uniquely determined by its actions on all the parameters:

$$\nu_u(a) = \chi_{F(a)}(u), \forall a \in A,$$

where $\chi(\cdot)$ denotes the characteristic function of a set.

Given an object $u \in U$ and a formula $\varphi \in \mathscr{F}(S)$, we say that u satisfies the formula φ in the soft set $S = (F, A)$, denoted $u \vDash_S \varphi$, if $\nu_u(\varphi) = 1$. By $\|\varphi\|_S$ we denote the semantic interpretation (meaning) of the formula $\varphi \in \mathscr{F}(S)$, which is defined by

$$\|\varphi\|_S = \{u \in U : \nu_u(\varphi) = 1\}.$$

Then it follows that $u \vDash_S \varphi$ if and only if $u \in \|\varphi\|_S$. The following result is easily obtained from the above definitions.

Proposition 1. *Let $S = (F, A)$ be a soft set over U. We have the following:*

(1) $\|a\|_S = F(a)$, $\forall a \in A$;

(2) $\|\neg\varphi\|_S = U - \|\varphi\|_S$, $\forall \varphi \in \mathscr{F}(S)$;

(3) $\|\varphi \wedge \psi\|_S = \|\varphi\|_S \cap \|\psi\|_S$, $\forall \varphi, \psi \in \mathscr{F}(S)$;

(4) $\|\varphi \vee \psi\|_S = \|\varphi\|_S \cup \|\psi\|_S$, $\forall \varphi, \psi \in \mathscr{F}(S)$;

(5) $\|\varphi \to \psi\|_S = \|\neg\varphi \vee \psi\|_S = (U - \|\varphi\|_S) \cup \|\psi\|_S$, $\forall \varphi, \psi \in \mathscr{F}(S)$.

3.3 Evaluate the Soundness of Formulas

As was mentioned above, the notion of the degree of the truth of formulas plays a fundamental role in quantitative logic. In a similar fashion, we shall introduce below some useful notions to evaluate the soundness of formulas over a given soft set.

Definition 7. (Basic soft truth degree) *Let $S = (F, A)$ be a soft set over U and $\varphi \in \mathscr{F}(S)$. The* basic soft truth degree *of φ is defined by*

$$\beta_S(\varphi) = \frac{1}{|U|} \sum_{u \in U} \nu_u(\varphi) \tag{2}$$

where $|\cdot|$ denotes the cardinality of the set.

It is easy to see that $0 \le \beta_S(\varphi) \le 1$ for all $\varphi \in \mathscr{F}(S)$. Also we have $\beta_S(\varphi) = 0$ (resp. $\beta_S(\varphi) = 1$) if and only if $\|\varphi\|_S = \emptyset$ (resp. $\|\varphi\|_S = U$). Note also that both $\beta_S(\cdot)$ and $\tau_2(\cdot)$ can be viewed as "voting models". The truth degree is then expressed exactly by the support rate of the given formula.

The major difference between $\beta_S(\cdot)$ and $\tau_2(\cdot)$ lies in the fact that for $\beta_S(\cdot)$ referees are just objects in U whose "attitude" is specified by the soft set $S = (F, A)$, while for $\tau_2(\cdot)$ referees are all 0-1 vectors $(x_1, \cdots, x_m) \in \{0, 1\}^m$ without any further restriction. In other words, $\tau_2(\cdot)$ depends only on the structure of the propositional logic system; the basic soft truth degree $\beta_S(\cdot)$, however, depends also on the structure of the soft set. Hence for a given formula $\varphi \in \mathscr{F}(S')$, we may have $\beta_{S'}(\varphi) \neq \beta_S(\varphi)$ even if S' is a soft subset of S.

From the above discussion we know that soundness of a formula can not only be graded but also be viewed as a concept relative to the universe of discourse. This point is reasonable as we know that "theory of everything" is still a dream for us and we must restrict to a certain universe when apply some "known truth" in real-world applications.

3.4 Decision Rules

Maji et al. initiated the application of soft sets to decision making problems in [10]. It is worth noting that the soft set based decision making in the sense of Maji et al. [10] is actually a question concerning the selection of optimal alternatives. Almost all other research in this direction is restricted to the discussion of this issue.

The notion of *choice values* is of vital importance in coping with these decision making problems since choice values express the number of good attributes possessed by an object. It follows that the optimal decision is just to select the object with the maximum choice value.

On the other hand we should know that the most important thing for decision making (in a more general sense) is to induce useful decision rules from some "training samples". These decision rules will be used in decision support, inference, knowledge discovery, prediction and some related areas in artificial intelligence.

Let $S = (F, C, D)$ be a decision soft set over U. A *decision rule* in S is any formula of the form $\varphi \to \psi$, where $\varphi \in \mathscr{F}(S_C)$ and $\psi \in \mathscr{F}(S_D)$. The condition formula φ and the decision formula ψ are also referred to as the *predecessor* and the *successor* of the decision rule $\varphi \to \psi$.

A decision rule $\varphi \to \psi$ is *absolutely true* in S if $\|\varphi\|_S \subseteq \|\psi\|_S$. Since a decision rule is still a formula over a decision soft set $S = (F, C, D)$, we can also measure its truth degree by using the concept of basic soft truth degrees introduced above.

Proposition 2. *Given a decision soft set $S = (F, C, D)$ over U and a decision rule $\varphi \to \psi$ in S, we have that $\varphi \to \psi$ is absolutely true if and only if $\beta_S(\varphi \to \psi) = 1$.*

Proof. Assume that the decision rule $\varphi \to \psi$ is absolutely true in S, i.e., $\|\varphi\|_S \subseteq \|\psi\|_S$. By Proposition 1, we know that

$$\|\varphi \to \psi\|_S = \|\neg\varphi \vee \psi\|_S = (U - \|\varphi\|_S) \cup \|\psi\|_S.$$

But we also have $U - \|\varphi\|_S \supseteq U - \|\psi\|_S$ since $\|\varphi\|_S \subseteq \|\psi\|_S$. Hence $\|\varphi \to \psi\|_S = U$, and so we deduce that $\beta_S(\varphi \to \psi) = 1$.

Conversely, if $\beta_S(\varphi \to \psi) = 1$ then $\|\varphi \to \psi\|_S = U$. Now let $u \in \|\varphi\|_S$, i.e., $\nu_u(\varphi) = 1$. It follows that $\nu_u(\psi) = 1$ since we know that $\nu_u(\varphi \to \psi) = 1$ holds for all $u \in U$. This says that $\|\varphi\|_S \subseteq \|\psi\|_S$ as required. □

Notice that a decision rule $\varphi \to \psi$ in S with $\|\varphi\|_S = \emptyset$ is always absolutely true. This is valid in logical sense but not very reasonable for practical needs. In fact we usually take rules of this type as pseudo ones in practice. So one can sometimes measure the truth degree of a decision rule $\varphi \to \psi$ in S by introducing the following inclusion measure of $\|\varphi\|_S$ in $\|\psi\|_S$.

Definition 8. (Conditional soft truth degree) *Given a decision soft set $S = (F, C, D)$ over U and a decision rule $\varphi \to \psi$ in S with $\beta_S(\varphi) \neq 0$. The* conditional soft truth degree *of $\varphi \to \psi$ is defined by*

$$\gamma_S(\varphi \to \psi) = \frac{\beta_S(\varphi \wedge \psi)}{\beta_S(\varphi)}. \tag{3}$$

Proposition 3. *Given a decision soft set $S = (F, C, D)$ over U and a decision rule $\varphi \to \psi$ in S with $\beta_S(\varphi) \neq 0$. We have that $\varphi \to \psi$ is absolutely true if and only if $\gamma_S(\varphi \to \psi) = 1$.*

Proof. By Proposition 1, we have $\|\varphi \wedge \psi\|_S = \|\varphi\|_S \cap \|\psi\|_S$. Thus

$$\gamma_S(\varphi \to \psi) = 1 \Leftrightarrow \|\varphi\|_S \cap \|\psi\|_S = \|\varphi\|_S \Leftrightarrow \|\varphi\|_S \subseteq \|\psi\|_S.$$

This completes the proof. □

The conditional soft truth degree defined above can also be interpreted as the conditional probability that the successor ψ is true given the predecessor φ. As an alternative way to evaluate the soundness of decision rules we may use the following measure.

Definition 9. (Product soft truth degree) *Given a decision soft set $S = (F, C, D)$ over U and a decision rule $\varphi \to \psi$ in S. The* product soft truth degree *of $\varphi \to \psi$ is defined by*

$$\varrho_S(\varphi \to \psi) = \beta_S(\varphi \to \psi) \cdot \beta_S(\varphi \wedge \psi). \tag{4}$$

It should be noted that the product soft truth degree is a more strict measure for evaluating the soundness of decision rules since it can be used to exclude those pseudo absolutely true rules. Specifically, for any decision rule $\varphi \to \psi$ in S with $\|\varphi\|_S = \emptyset$, we have $\beta_S(\varphi \wedge \psi) = 0$ even if $\beta_S(\varphi \to \psi) = 1$; hence finally we obtain that $\varrho_S(\varphi \to \psi) = 0$ in this case.

Proposition 4. *Given a decision soft set $S = (F, C, D)$ over U and a decision rule $\varphi \to \psi$ in S. If $\varphi \to \psi$ is absolutely true, then we have $\varrho_S(\varphi \to \psi) = \beta_S(\varphi)$.*

Proof. Assume that the decision rule $\varphi \to \psi$ is absolutely true in S, i.e., $\|\varphi\|_S \subseteq \|\psi\|_S$. By Proposition 1, $\|\varphi \wedge \psi\|_S = \|\varphi\|_S \cap \|\psi\|_S$. Thus we deduce that $\|\varphi \wedge \psi\|_S = \|\varphi\|_S$, whence $\beta_S(\varphi \wedge \psi) = \beta_S(\varphi)$. Note also that $\beta_S(\varphi \to \psi) = 1$ by Proposition 2. Hence it follows that

$$\varrho_S(\varphi \to \psi) = 1 \cdot \beta_S(\varphi).$$

This completes the proof.

3.5 Rule Evaluation

After rule induction we need to compare different rules extracted from the data of the given soft set so as to find the most useful decision rules. Next, we shall introduce some preorders for comparing decision rules.

Definition 10. (Basic rule preorder) *Given a decision soft set $S = (F, C, D)$ over U. For decision rules $\varphi \to \psi$ and $\varphi' \to \psi'$ in S, we define*

$$\varphi \to \psi \preceq_\beta \varphi' \to \psi' \Leftrightarrow (\beta_S(\varphi \to \psi), \beta_S(\varphi)) \leqslant_l (\beta_S(\varphi' \to \psi'), \beta_S(\varphi')),$$

called the basic rule preorder *on S. Here $\leqslant_l$ denotes the lexicographical order on I^2.*

We say that the decision rule $\varphi' \to \psi'$ is better than $\varphi \to \psi$ (with respect to the basic rule preorder $\preceq_\beta$) if $\varphi \to \psi \preceq_\beta \varphi' \to \psi'$ holds.

Definition 11. (Product rule preorder) *Let $S = (F, C, D)$ be a decision soft set over U. For decision rules $\varphi \to \psi$ and $\varphi' \to \psi'$ in S, we define*

$$\varphi \to \psi \preceq_\varrho \varphi' \to \psi' \Leftrightarrow (\varrho_S(\varphi \to \psi), \beta_S(\varphi)) \leqslant_l (\varrho_S(\varphi' \to \psi'), \beta_S(\varphi')),$$

called the product rule preorder *on S.*

Note that the *conditional rule preorder* $\preceq_\gamma$ can also be defined in an expected way for rules $\varphi \to \psi$ and $\varphi' \to \psi'$ such that $\beta_S(\varphi) \neq 0$ and $\beta_S(\varphi') \neq 0$. By the above preorders, decision rules are estimated according to two aspects; one concerns the degree of applicability of the rule, and the other is associated with the soundness of the rule.

Definition 12. (Atomic rule) *Let $S = (F, C, D)$ be a decision soft set over U. A decision rule $\varphi \to \psi$ in S is called an* atomic rule *if $\varphi \in C$ and $\psi \in D$. The set of all atomic rules in the soft set S is denoted by $\mathscr{R}_A(S)$.*

We say that an atomic rule $\varphi \to \psi$ is *B-optimal* (resp. *C-optimal, P-optimal*) if it is maximal in $\mathscr{R}_A(S)$ with respect to the rule preorder $\preceq_\beta$ (resp. $\preceq_\gamma$, $\preceq_\varrho$).

4 An Illustrative Example

Here we present an interesting example to illustrate all the new concepts initiated in this study.

Suppose that there are six houses under our consideration, namely the universe $U = \{h_1, h_2, h_3, h_4, h_5, h_6\}$, and the condition parameter set $C = \{e_1, e_2, e_3, e_4\}$, where e_i respectively stand for "beautiful ", "modern", "cheap" and "in green surroundings". Assume that the decision parameter set $D = \{d\}$, where d means "attractive". All the information available on these houses can be characterized by a decision soft set $S = (F, C, D)$, with its tabular representation shown in Table 1.

Table 1. A decision soft set $S = (F, C, D)$

House	e_1	e_2	e_3	e_4	d
h_1	1	1	1	1	1
h_2	0	0	1	0	0
h_3	1	1	0	0	0
h_4	1	0	1	1	1
h_5	0	0	1	0	1
h_6	1	0	0	1	1

It is easy to see the semantic meaning of an atomic formula. For instance, we have $\|e_1\|_S = F(e_1)$, which says that the set of "beautiful houses" is $\{h_1, h_3, h_4, h_6\}$. Similarly, we can see from $\|e_1 \wedge e_3\|_S = F(e_1) \cap F(e_3)$ that h_1 and h_4 are "beautiful and cheap houses". In general, $\|\varphi\|_S$ will define the meaning of the formula φ in the setting of the soft set S.

Now we consider how to induce useful decision rules in S. For the sake of convenience, we shall only consider some "simple" decision rules, particularly, the atomic rules so that we could concentrate on illustrating our basic ideas. First we calculate the basic soft truth degrees of the atomic rules. We obtain that $\beta_S(e_i \to d) = 5/6$ $(i = 1, 2, 3)$ and $\beta_S(e_4 \to d) = 1$. For conditional soft truth degrees of these atomic rules, we have $\gamma_S(e_i \to d) = 3/4$ $(i = 1, 3)$, $\gamma_S(e_2 \to d) = 1/2$ and $\gamma_S(e_4 \to d) = 1$. Also we can see $\beta_S(e_i \wedge d) = 1/2$ $(i = 1, 3)$, $\beta_S(e_2 \wedge d) = 1/6$ and $\beta_S(e_4 \wedge d) = 1/2$. Thus by easy calculation we know the product soft truth degrees of these atomic rules, namely $\varrho_S(e_i \to d) = 5/12$ $(i = 1, 3)$, $\varrho_S(e_2 \to d) = 5/36$ and $\varrho_S(e_4 \to d) = \beta_S(e_4 \wedge d) = 1/2$.

Next let us compare the above atomic rules by using the rule preorders. Note that $\beta_S(e_i) = 2/3$ $(i = 1, 3)$, $\beta_S(e_2) = 1/3$ and $\beta_S(e_4) = 1/2$. Thus $e_4 \to d$ is the B-optimal, C-optimal and P-optimal decision rule. We can conclude that "in green surroundings" is the most important aspect to determine whether a house is "attractive", according to what we have learned from the collected data.

5 Conclusion

We have established in this study a preliminary version of the theory of approximate reasoning based on soft sets. The basic ideas of the proposed scheme was motivated by quantitative logic. We also contributed to soft set based decision making by introducing decision rules in soft sets and focusing on inducing useful rules in decision making. As future work, connections between soft sets and non-classical logic could be explored.

Acknowledgements. This work was partially supported by the National Natural Science Foundation of China (Nos.10571112, 60873119 and 10926031), the Higher School Doctoral Subject Foundation of Ministry of Education of China (No. 200807180005), the Shaanxi Provincial Research and Development Plan of Science and Technology (No. 2008K01-33), the Natural Science Foundation of Educational Committee of Shaanxi Province of China (Nos. 09JK722, 09JK724 and 09JK728), and the Research Grant of Xi'an Institute of Posts and Telecommunications (No. ZL2010-30). The research of F. Feng is also supported by a grant from the Education Department of Shaanxi Province of China ("Soft set theory and its applications in artificial intelligence with uncertainty", approved in 2010).

References

1. Zadeh, L.A.: Fuzzy sets. Information and Control 8, 338–353 (1965)
2. Pawlak, Z.: Rough sets. International Journal of Computer and Information Sciences 11, 341–356 (1982)
3. Gorzalzany, M.B.: A method of inference in approximate reasoning based on interval-valued fuzzy sets. Fuzzy Sets and Systems 21, 1–17 (1987)
4. Molodtsov, D.: Soft set theory–First results. Computers and Mathematics with Applications 37, 19–31 (1999)
5. Molodtsov, D.: The theory of soft sets. URSS Publishers, Moscow (2004)
6. Wang, G.J., Zhou, H.J.: Quantitative logic. Information Sciences 179, 226–247 (2009)
7. Maji, P.K., Biswas, R., Roy, A.R.: Soft set theory. Computers and Mathematics with Applications 45, 555–562 (2003)
8. Feng, F., Li, C.X., Davvaz, B., Ali, M.I.: Soft sets combined with fuzzy sets and rough sets: a tentative approach. Soft Computing 14, 899–911 (2010)
9. Pawlak, Z., Skowron, A.: Rudiments of rough sets. Information Sciences 177, 3–27 (2007)
10. Maji, P.K., Roy, A.R., Biswas, R.: An application of soft sets in a decision making problem. Computers and Mathematics with Applications 44, 1077–1083 (2002)

Left(Right)-Continuity of t-Norms on the Metric Lattice

Li-na Ma[1], Shuo Liu[2], and Guo-jun Wang[1]

[1] College of Mathematics and Information Science, Shaanxi Normal University, Xi'an 710062, China
e-mail: {malina,gjwang}@snnu.edu.cn

[2] Faculty of Biomedical Engineering, The Fourth Military Medical University, Xi'an 710032, China
e-mail: liushuo912@163.com

Abstract. Left-continuity of t-norms on the unit interval $[0,1]$ is equivalent to the property of sup-preserving, but this equivalence does not hold for t-norms on the general complete lattice. In this paper, we initially introduce a special kind of the complete lattice–metric lattice. Based on the theory of directed partial order set, we prove that a t-norm on the metric lattice is left(right)-continuous if and only if it preserves direct sups(infs).

Keywords: T-norm, metric lattice, left(Right)-continuity, direct-sup(inf)-preserving.

1 Introduction

An important notion in fuzzy set theory is that of triangular norms which were introduced by Schweizer and Sklar in the framework of probabilistic metric spaces[1]. The condition of their left-continuity is crucial in almost all such fields, in particular in fuzzy logic, where this property makes it possible to evaluate the implication by the residuum of the conjunction. On the unit interval, the "sup-preserving" is used to describe the "left-continuity" of t-norm, that is, a left-continuous t-norm $\otimes$ on the unit interval $[0,1]$ is equivalent to the property of sup-preserving(or $\otimes$ satisfies the residuation principle[2]), but this is no longer true for t-norms on the general complete lattice and a strict counter-example was constructed in [3]. Does there exists any interrelations between left-continuity and sup-preserving property for t-norms on the certain kind of complete lattice? The present paper is to give a positive answer to this question. we initially introduce a special kind of complete lattice–metric lattice, then we prove that a t-norm $\otimes$ on a metric lattice is left-continuous if and only if $\otimes$ preserves direct sups, and $\otimes$ is right-continuous if and only if $\otimes$ preserves direct infs.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 257–261.
springerlink.com

2 Preliminaries

Definition 1. ([5]) *Let $(L, \leq)$ be a partial order set, D be a non-empty subset of L. If $\forall a, b \in D$, there exists $c \in D$ such that $a \leq c$ and $b \leq c$, then D is said to be an upper directed set; if exists $c \in D$ such that $c \leq a$ and $c \leq b$, then D is said to be a lower directed set.*

Definition 2. ([5]) *Let D be a directed set, L be a non-empty set, the mapping $S : D \longmapsto X$ is said to be the net of L.*

Specially, let L be a complete lattice, then we give the concept of upper(lower) limit points(or convergence points) of a net of L as follows:

Definition 3. *Let L be a complete lattice, ρ be a metric on L, $a \in L$ and $S : D \longmapsto L$ be a net,*

(i) suppose that D is an upper directed set, if $\forall \varepsilon > 0$, there exists $n_0 \in D$ such that $\rho(n, a) < \varepsilon$ whenever $n \in D$ and $n \geq n_0$, then we say S is upper limited to a(or S is upper convergent to a);

(ii) suppose that D is a lower directed set, if $\forall \varepsilon > 0$, there exists $n_0 \in D$ such that $\rho(n, a) < \varepsilon$ whenever $n \in D$ and $n \leq n_0$, then we say S is lower limited to a(or S is lower convergent to a).

Then we introduce the definition of a metric lattice.

Definition 4. *Let L be a complete lattice, $'$ be an order reversing involution on L, ρ be a metric on L such that*

(i) $max\{\rho(b, c), \rho(a, b)\} \leq \rho(a, c)$, whenever $a \leq b \leq c$;

(ii) Let $S : D \longmapsto L$ be a net where D is an upper directed set on L and $SupD = a(a \in L)$, then S be upper convergent to a.

Then L is said to be a metric lattice.

Example 1. (i) Let $L = [0, 1]^2, \rho$ be an Euclidean metric on L, $(a^{'}, b^{'}) = (1 - a, 1 - b)$ where $(a, b) \in [0, 1]^2$, then L is a metric lattice.

(ii) Let $L = [0, \frac{1}{2}] \cup \{1\}$, ρ and $'$ are the same to (i), then L is not a metric lattice. In fact, it can be easily checked that L is an upper directed set and $SupL = 1$. Let $S : L \longmapsto L$ is an identity map on L, then S be a net of L, however S be not upper convergent to 1.

In the following, we assume that L is a metric lattice, $'$ be an order reversing involution, and 1, 0 are the largest element and the least element of L, respectively.

Definition 5. *([1,6])A triangular norm T (briefly t-norm)on L is a binary operator which is commutative, associative, monotone and has the neutral element 1.*

For the sake of convenience, we use $a \otimes b$ instead of $T(a, b)$, then $\otimes$ is a t-norm on L if the following conditions are satisfied:

(i) $a \otimes b = b \otimes a$;
(ii) $(a \otimes b) \otimes c = a \otimes (b \otimes c)$;
(iii) *if* $b \leq c$, *then* $a \otimes b \leq a \otimes c$;
(iv) $a \otimes 1 = a$,
where $a, b, c \in L$.

Definition 6. *([3]) Let* $\otimes$ *be a t-norm on* L.
(i) If for every non-empty subset A *of* L,

$$a \otimes sup\{z|z \in A\} = sup\{a \otimes z|z \in A\}, a \in L, \tag{1}$$

then we say $\otimes$ *is sup-preserving, or* $\otimes$ *preserves sups;*
(ii) If for every non-empty subset A *of* L,

$$a \otimes inf\{z|z \in A\} = inf\{a \otimes z|z \in A\}, a \in L, \tag{2}$$

then we say $\otimes$ *is inf-preserving, or* $\otimes$ *preserves infs;*
(iii) If the equation (1)((2)) holds for upper(lower)directed set A, *then we say that* $\otimes$ *preserves direct sups(infs) (briefly,* $\otimes$ *preserves dsups(dinfs)).*

Definition 7. *Let* $\otimes$ *be a t-norm on* L, *then* $D = \{x \in L|x \leq \beta, \beta \in L\}$ *and* $D^0 = \{a \otimes x|x \in D, a \in L\}$ *are two upper directed sets of* L, *hence there are two nets of* L, *one is* $S : D \longmapsto L$ *and the other is* $S^0 : D^0 \longmapsto L$. *If we have* S^0 *is upper(lower) limited to* $a \otimes \beta$ *whenever* S *being upper(lower) convergent to* β, *then* $\otimes$ *is called left(right)-continuous.*

3 A Necessary and Sufficient Condition for t-Norms on L Being Left(right)-Continuous

The following theorem show that left-continuous is equivalent to dsup-preserving for t-norms on the metric lattice.

Theorem 1. *Let* $\otimes$ *be a t-norm on* L, *then* $\otimes$ *is left-continuous if and only if* $\otimes$ *preserves dsups.*

Proof. Suppose that $\otimes$ is left-continuous, D is an upper directed subset of $L, S : D \longmapsto L$ be a net, $\beta = supD$ and

$$a \otimes \beta \neq sup\{a \otimes z|z \in D\}. \tag{3}$$

Let

$$\alpha = sup\{a \otimes z|z \in D\}. \tag{4}$$

By the monotonicity of $\otimes$ and $\beta = supD$, we have $\alpha < a \otimes \beta$. Hence

$$\rho(\alpha, a \otimes \beta) = \delta > 0. \tag{5}$$

Assume that $x, y \in D$, then there exists $z \in D$ such that $x \leq z$ and $y \leq z$, hence $a \otimes x \leq a \otimes z$ and $a \otimes y \leq a \otimes z$, and $\{a \otimes z|z \in D\}$ is an upper directed subset of L. Let $D^0 = \{a \otimes z | z \in D\}$, then it follows from (4) that the net $S^0 : D^0 \longmapsto L$ be upper convergent to α, hence for $\varepsilon = \frac{\delta}{2} > 0$, there exists $a \otimes z_0 \in D^0, z_0 \in D$ such that

$$\rho(a \otimes z_1, \alpha) < \frac{\delta}{2}.$$

when $a \otimes z_1 \in D^0$ and $a \otimes z_1 \geq a \otimes z_0$.

Owning to $supD = \beta$, S is convergent to β. By the left-continuity of $\otimes$ it follows that S^0 is limited to $a \otimes \beta$. Then for $\varepsilon = \frac{\delta}{2} > 0$, there exists $a \otimes y_0 \in D^0$ such that

$$\rho(a \otimes y_1, a \otimes \beta) < \frac{\delta}{2}$$

where $a \otimes y_1 \in D^0$ and $a \otimes y_1 \geq a \otimes y_0$.

Since D is upper directed there exists $\overline{x} \in D$ such that $z_1 \leq \overline{x}, y_1 \leq \overline{x}$. Then it follows from (4) that $a \otimes z_1 \leq a \otimes \overline{x} \leq \alpha$, hence we obtain

$$\rho(a \otimes \overline{x}, \alpha) \leq \rho(a \otimes z_1, \alpha) < \frac{\delta}{2}.$$

Moreover, since $\overline{x} \in D, \beta = supD$, we obtain $\overline{x} \leq \beta$ and $a \otimes y_1 \leq a \otimes \overline{x} \leq a \otimes \beta$, then we have

$$\rho(a \otimes \overline{x}, a \otimes \beta) \leq \rho(a \otimes y_1, a \otimes \beta) < \frac{\delta}{2}.$$

Therefore

$$\rho(\alpha, a \otimes \beta) \leq \rho(\alpha, a \otimes \overline{x}) + \rho(a \otimes \overline{x}, a \otimes \beta) < \frac{\delta}{2} + \frac{\delta}{2} = \delta.$$

This contradicts (5). Hence (3) does not hold and $\otimes$ is dsup-preserving.

Conversely, suppose that $\otimes$ is dsup-preserving on L, we are to prove that $\otimes$ is left-continuous on L. Let

$$D = \{x \in L | x \leq \beta, \beta \in L\}, D^0 = \{a \otimes x | x \in D, a \in L\}$$

then $S : D \longmapsto L$ and $S^0 : D^0 \longmapsto L$ are two nets of L. we only need to prove that S^0 is upper limited to $a \otimes \beta$ whenever S being upper convergent to β.

It is clear that β is an upper bound of D. Suppose that α is the other upper bound of D, then $\beta \leq \alpha$ because $\beta \in D$, that is, $supD = \beta$. Since $\otimes$ is dsup-preserving it follows that

$$a \otimes \beta = a \otimes sup\{x \in L | x \leq \beta\} = sup\{a \otimes x | x \in D\} = supD^0.$$

Then D^0 be upper convergent to $a \otimes \beta$. This proves that $\otimes$ is left-continuous on L.

Note that "$\prime$" be an order reversing involution on L, then we can prove the following corollary which gives the necessary and sufficient condition of a t-norm $\otimes$ on the metric lattice being right-continuous by the completely dual approach.

Corollary 1. *Let $\otimes$ be a t-norm on L, then $\otimes$ is right-continuous if and only if $\otimes$ preserves dinfs.*

4 Conclusion

In the present paper, the concept of metric lattice are initially introduced. Then, based on the theory of directed partial order set we give a necessary and sufficient condition for t-norms on metric lattice being left(right)-continuous.

References

1. Schweizer, B., Sklar, A.: Probability metric space. North-Holland, New york (1983)
2. Wang, G.J.: Introduction to Mathematical and Resolution Principle. Science Press, Beijing (2008)
3. Wang, G.J., Wang, W.: Left-continuity of t-norms on the n-dimensional Euclidean cube. Applied Mathematics Letters 23(4), 479–483 (2010)
4. Song, J., Luo, M.N., Li, Y.M.: Necessary and suffient condition for left-continuity of interval-valued triangular norms. Journal of Shangdong University (Natural Science) 44(3), 67–70 (2009)
5. Wang, G.J.: Theory of L-fuzzy topological spaces. Shaanxi Normal University Press (1988)
6. Klement, E.P., Mesiar, R., Pap, E.: Triangular norms. Kluwer Academic Publishers, Dordrent (2000)
7. Deschrijver, G.: A representation of t-norms in interva-valued L-fuzzy set theory. Fuzzy Set and Systems 159(13), 1593–1618 (2008)
8. Deschrijver, G., Cornelis, C., Kerre, E.E.: On the representation of intuitionistic fuzzy t-norms and t-conorms. IEEE Transactions on Fuzzy Systems 12(1), 45–61 (2004)
9. Deschrijver, G.: Characterization of (weekly) Archimedean t-norms in interval-valued fuzzy set theory. Fuzzy Set and Systems 160(6), 778–801 (2009)

Predicate Formal System $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$ and Its Completeness

Ying-cang Ma and Huan Liu

School of Science, Xi'an Polytechnic University, Xi'an 710048, China
e-mail: mayingcang@126.com

Abstract. The propositional calculus formal deductive system $\mathcal{UL}^{\triangle}_{h\in(0,1]}$ for 0-level universal AND operator with projection operator has been built up. In this paper, according to the propositional system, a predicate calculus formal deductive system $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$ for 0-level universal AND operator with projection operator is built up. The completeness theorem of system $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$ are given. So it shows that the Semantic and Syntactic of system $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$ are harmony.

Keywords: Universal logic, fuzzy logic, predicate calculus, axiomatization.

1 Introduction

How to deal with various uncertainties and evolution problems have been critical issues for further development of AI. The well-developed mathematical logic is too rigid and it can only solve certainty problems. It is the new challenge for logics to make mathematical logic more flexible and to contain various uncertainties and evolution. Therefore, non-classical logic and modern logic develop rapidly, for example fuzzy logic and universal logic.

In recent years considerable progress has been made in logical foundations of fuzzy logic, especially for the logic based on t-norm and its residua (See [1-11]). Some well-known logic systems have been built up, such as the basic logic (BL)[1,3] introduced by Hajek; the monoidal t-norm based logic (MTL)[2] introduced by Esteva and Godo; a formal deductive system L^* introduced by Wang (see [7-11]), and so on. Moreover the completeness of the above logical systems have been proven.

Universal logic[12] was proposed by Huacan He, which thinks that all things in the world are correlative, that is, they are either mutually exclusive or mutually consistent, and we call this kind of relation generalized correlation. Any two propositions have generalized correlation. The degree of general correlation can be described quantitatively by the coefficient of the generalized

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 263–272.
springerlink.com

correlation $h \in [0,1]$: If we define the h of operator $T(a,b)$ as the ratio between the volume of $T(a,b)$ and the volume of maximal operator, then $h = 1$ means the maximal attractive state; $h = 0.75$ means independency correlative state; $h = 0.5$ means neutral state; $h = 0$ means maximal exclusive state. The 0-level universal AND operators and 0-level universal IMPLICATION operators are defined as:

0-level universal AND operators are mapping $T : [0,1] \times [0,1] \to [0,1]$, $T(x,y,h) = \Gamma^1[(x^m + y^m - 1)^{1/m}]$, which is usually denoted by $\wedge_h$; the real number m has relation with the coefficient of generalized correlation h as:

$$m = (3 - 4h)/(4h(1-h)) \tag{1}$$

$h \in [0,1], m \in \mathbb{R}$. And $\Gamma^1[x]$ denotes x is restricted in $[0,1]$, if $x > 1$ then its value will be 1, if $x < 0$, its value will be 0.

0-level universal IMPLICATION operators are mapping $I : [0,1] \times [0,1] \to [0,1], I(x,y,h) = ite\{1|x \leq y; 0|m \leq 0 \text{ and } y = 0; \Gamma^1[(1 - x^m + y^m)^{1/m}]\}$, which is usually denoted by $\Rightarrow_h$. In the above the equation with m and h is the same as (1).

The formal systems of propositional universal logic have been studied in [13-18]. In [18], the soundness of predicate calculus formal deductive systems $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$ has been studied. In this paper, we focus on the completeness of predicate formal system $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$.

The paper is organized as follows. After this introduction, Section 2 we will give the predicate calculus formal deductive system $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$ and introduce its soundness theorem. In Section 3 the completeness of system $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$ will be proved. Some extension logic systems of $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$ are introduced in Section 4. The final section offers the conclusion.

2 Predicate Formal System $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$

In this section, we will introduce some basis definition and important results of System $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$ have obtained in [18],

First-order language J consists of symbols set and generation rules:

The symbols set of J consist of as following:

(1) Object variables: $x, y, z, x_1, y_1, z_1, x_2, y_2, z_2, \cdots$;

(2) Object constants: a, b, c, a_1, b_1, c_1,, Truth constants: $\overline{0}, \overline{1}$;

(3) Predicate symbols: $P, Q, R, P_1, Q_1, R_1, \cdots$;

(4) Connectives: $\&, \to, \triangle$;

(5) Quantifiers: $\forall$(universal quantifier), $\exists$(existential quantifier);

(6) Auxiliary symbols: (,),,.

The symbols in (1)-(3) are called non-logical symbols of language J. The object variables and object constants of J are called terms. The set of all object constants is denoted by Var(J), The set of all object variables is denoted by Const(J), The set of all terms is denoted by Term(J). If P is n-ary predicate symbol, $t_1, t_2, \cdots, t_n$ are terms, then $P(t_1, t_2, \cdots, t_n)$ is called atomic formula.

The formula set of J is generated by the following three rules in finite times:

(i) If P is atomic formula, then $P \in J$;

(ii) If $P, Q \in J$, then $P\&Q, P \to Q, \triangle P \in J$;

(iii) If $P \in J$, and $x \in \mathrm{Var}(J)$, then $(\forall x)P, (\exists x)P \in J$.

The formulas of J can be denoted by $\varphi, \phi, \psi, \varphi_1, \phi_1, \psi_1, \cdots$. Further connectives are defined as following:

$\varphi \wedge \psi$ is $\varphi\&(\varphi \to \psi)$, $\varphi \vee \psi$ is $((\varphi \to \psi) \to \psi) \wedge (\psi \to \varphi) \to \varphi)$,
$\neg\varphi$ is $\varphi \to \overline{0}$, $\varphi \equiv \psi$ is $(\varphi \to \psi)\&(\psi \to \varphi)$.

Definition 1. *The axioms and deduction rules of predicate formal system $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$ as following:*

(i)The following formulas are axioms of $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$:

(U1) $(\varphi \to \psi) \to ((\psi \to \chi)(\varphi \to \chi))$
(U2) $(\varphi\&\psi) \to \varphi$
(U3) $(\varphi\&\psi) \to (\psi\&\varphi)$
(U4) $\varphi\&(\varphi \to \psi) \to (\psi\&(\psi \to \varphi))$
(U5) $(\varphi \to (\psi \to \chi)) \to ((\varphi\&\psi) \to \chi)$
(U6) $((\varphi\&\psi) \to \chi) \to (\varphi \to (\psi \to \chi))$
(U7) $((\varphi \to \psi) \to \chi) \to (((\psi \to \varphi) \to \chi) \to \chi)$
(U8) $\overline{0} \to \varphi$
(U9) $(\varphi \to \varphi\&\psi) \to ((\varphi \to \overline{0}) \vee \psi \vee ((\varphi \to \varphi\&\varphi) \wedge (\psi \to \psi\&\psi)))$
(U10) $\triangle\varphi \vee \neg\triangle\varphi$
(U11) $\triangle(\varphi \vee \psi) \to (\triangle\varphi \vee \triangle\psi)$
(U12) $\triangle\varphi \to \varphi$
(U13) $\triangle\varphi \to \triangle\triangle\varphi$
(U14) $\triangle(\varphi \to \psi) \to (\triangle\varphi \to \triangle\psi)$
(U15) $(\forall x)\varphi(x) \to \varphi(t)$ *(t substitutable for x in $\varphi(x)$)*
(U16) $\varphi(t) \to (\exists x)\varphi(x)$*($t$ substitutable for x in $\varphi(x)$)*
(U17) $(\forall x)(\chi \to \varphi) \to (\chi \to (\forall x)\varphi)$*($x$ is not free in χ)*
(U18) $(\forall x)(\varphi \to \chi) \to ((\exists x)\varphi \to \chi)$*($x$ is not free in χ)*
(U19) $(\forall x)(\varphi \vee \chi) \to ((\forall x)\varphi \vee \chi)$*($x$ is not free in χ)*

Deduction rules of $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$ are: Modus Ponens(MP):from $\varphi, \varphi \to \psi$ infer ψ; Necessitation: from φ infer $\triangle\varphi$; Generalization: from φ infer $(\forall x)\varphi$.

The meaning of "t substitutable for x in $\varphi(x)$" and "x is not free in χ" in the above definition have the same meaning in the classical first-order predicate logic, moreover, we can define the concepts such as proof, theorem, theory, deduction from a theory T, T-consequence in the system $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$. $T \vdash \varphi$ denotes that φ is provable in the theory T. $\vdash \varphi$ denotes that φ is a theorem of system $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$. Let $\mathrm{Thm}(\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}) = \{\varphi \in J | \vdash \varphi\}, \mathrm{Ded}(T) = \{\varphi \in$

$J|T \vdash \varphi\}$. Being the axioms of propositional system $\mathcal{UL}^{\triangle}_{h\in(0,1]}$ are in predicate system $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$, then the theorems in $\mathcal{UL}_{h\in(0,1]}$ are theorems in $\forall\mathcal{UL}_{h\in(0,1]}$. According the similar proof in [1,15,16] we can get the following lemmas.

Lemma 1. *The hypothetical syllogism holds in* $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$, *i.e. let* $\Gamma = \{\varphi \to \psi, \psi \to \chi\}$, *then* $\Gamma \vdash \varphi \to \chi$.

Lemma 2. $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$ *proves:*
(1) $\varphi \to \varphi$; *(2)* $\varphi \to (\psi \to \varphi)$;
(3) $(\varphi \to \psi) \to ((\varphi \to \gamma) \to (\psi \to \gamma))$;
(4)$(\varphi\&(\varphi \to \psi)) \to \psi$; *(5)* $\Delta\varphi \equiv \Delta\varphi\&\Delta\varphi$.

Lemma 3. *If* $T = \{\varphi \to \psi, \chi \to \gamma\}$, *then* $T \vdash (\varphi\&\chi) \to (\psi\&\gamma)$.

In order to prove the soundness of predicate system $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$, we should introduce the following definitions.

Definition 2 *([1]). A BL-algebra is an algebra* $\mathrm{L} = (L, \cap, \cup, *, \Rightarrow, 0, 1)$ *with four binary operations and two constants such that*

1. $(L, \cap, \cup, 0, 1)$ *is a lattice with the greatest element* 1 *and the least element* 0 *(with respect to the lattice ordering* $\leq$*),*
2. $(L, *, 1)$ *is a commutative semigroup with the unit element* 1, *i.e.* $*$ *is commutative, associative and* $1 * x = x$ *for all* x,
3. *the conditions (i)* $z \leq (x \Rightarrow y)$ *iff* $x * z \leq y$; *(ii)* $x \cap y = x * (x \Rightarrow y)$; *(iii)* $(x \Rightarrow y) \cup (y \Rightarrow x) = 1$ *hold for all* x, y, z.

Definition 3 *([16]). A* ŁΠG *algebra is a BL-algebra in which the identity* $(x \Rightarrow x * y) \Rightarrow ((x \Rightarrow 0) \cup y \cup ((x \Rightarrow x * x) \cap (y \Rightarrow y * y))) = 1$ *is valid.*

Definition 4 *([15]). A* ŁΠG$_\triangle$*-algebra is a structure* $\mathrm{L} =< L, *, \Rightarrow, \cap, \cup, 0, 1, \triangle >$ *which is a* ŁΠG*-algebra expanded by an unary operation* $\triangle$ *in which the following formulas are true:*
$\triangle x \cup (\triangle x \Rightarrow 0) = 1$; $\triangle(x \cup y) \leq \triangle x \cup \triangle y$; $\triangle x \leq x$;
$\triangle x \leq \triangle\triangle x$; $(\triangle x) * (\triangle(x \Rightarrow y)) \leq \triangle y$; $\triangle 1 = 1$.

Let J is first-order predicate language, L is linearly ordered ŁΠG$_\triangle$ algebra, $\mathrm{M} = (M, (r_P)_P, (m_c)_c)$ is called a L-evaluation for first-order predicate language J, which M is non-empty domain, according to each n-ary predicate P and object constant c, r_P is L-fuzzy n-ary relation: $r_P : M^n \to \mathrm{L}$, m_c is an element of M.

Definition 5. *Let* J *be predicate language,* M *is* L*-evaluation of* J, x *is object variable,* $P \in J$.

(i) A mapping V*:* $\mathrm{Term}(J) \to M$ *is called* M*-evaluation, if for each* $c \in$Const (J), $v(c) = m_c$;

(ii)Two M*-evaluation* v, v' *are called equal denoted by* $v \equiv_x v'$ *if for each* $y \in \mathrm{Var}(J)\backslash\{x\}$, *there is* $v(y) = v'(y)$.

(iii) The value of a term given by M, v *is defined by:* $\|x\|_{\mathrm{M},v} = v(x)$; $\|c\|_{\mathrm{M},v} = m_c$. *We define the truth value* $\|\varphi\|^{\mathrm{L}}_{\mathrm{M},v}$ *of a formula* φ *as following. Clearly,* $*, \Rightarrow, \triangle$ *denote the operations of* L.

$\|P(t_1, t_2, \cdots, t_n)\|^{\mathrm{L}}_{\mathrm{M},v} = r_P(\|t_1\|_{\mathrm{M},v}, \cdots, \|t_n\|_{\mathrm{M},v})$

$\|\varphi \to \psi\|^{\mathrm{L}}_{\mathrm{M},v} = \|\varphi\|^{\mathrm{L}}_{\mathrm{M},v} \Rightarrow \|\psi\|^{\mathrm{L}}_{\mathrm{M},v}$

$\|\varphi \& \psi\|^{\mathrm{L}}_{\mathrm{M},v} = \|\varphi\|^{\mathrm{L}}_{\mathrm{M},v} * \|\psi\|^{\mathrm{L}}_{\mathrm{M},v}$

$\|\overline{0}\|^{\mathrm{L}}_{\mathrm{M},v} = 0; \quad \|\overline{1}\|^{\mathrm{L}}_{\mathrm{M},v} = 1$

$\|\Delta\varphi\|^{\mathrm{L}}_{\mathrm{M},v} = \Delta \|\varphi\|^{\mathrm{L}}_{\mathrm{M},v}$

$\|(\forall x)\varphi\|^{\mathrm{L}}_{\mathrm{M},v} = \inf\{\|\varphi\|^{\mathrm{L}}_{\mathrm{M},v'} \mid v \equiv_x v'\}$

$\|(\exists x)\varphi\|^{\mathrm{L}}_{\mathrm{M},v} = \sup\{\|\varphi\|^{\mathrm{L}}_{\mathrm{M},v'} \mid v \equiv_x v'\}$

In order to the above definitions are reasonable, the infimum/supremum should exist in the sense of L. So the structure M is L-safe if all the needed infima and suprema exist, i.e. $\|\varphi\|^{\mathrm{L}}_{\mathrm{M},v}$ is defined for all φ, v.

Definition 6. *Let* $\varphi \in J$, M *be a safe* L*-structure for* J.

(i) The truth value of φ *in* M *is* $\|\varphi\|^{\mathrm{L}}_{\mathrm{M}} = \inf\{\|\varphi\|^{\mathrm{L}}_{\mathrm{M},v} \mid v \ \ \mathrm{M}-\mathrm{evaluation}\}$.

(ii) A formula φ *of a language* J *is an* L*-tautology if* $\|\varphi\|^{\mathrm{L}}_{\mathrm{M}} = 1_{\mathrm{L}}$ *for each safe* L*-structure M. i.e.* $\|\varphi\|^{\mathrm{L}}_{\mathrm{M,v}} = 1$ *for each safe* L*-structure* M *and each* M*-valuation of object variables.*

Remark 1. For each $h \in (0,1], ([0,1], \wedge_h, \Rightarrow_h, \min, \max, 0, 1, \triangle)$ is a linearly ordered ŁΠG$_\triangle$-algebra. So the predicate system $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$ can be considered the axiomatization for 0-level universal AND operator with projection operator.

Definition 7. *Let* T *be a theory,* L *be a linearly ordered ŁΠG$_\triangle$-algebra and* M *a safe* L*-structure for the language of* T. M *is an* L*-model of* T *if all axioms of* T *are* $1_{\mathrm{L}}-$*true in* M*, i.e.* $\|\varphi\| = 1_L$ *in each* $\varphi \in T$.

Definition 8. *Let* T *be a theory,* L *be a linearly ordered ŁΠG$_\triangle$-algebra and* M *a safe* L*-structure for the language of* T. M *is an* L*-model of* T *if all axioms of* T *are* $1_{\mathrm{L}}-$*true in* M*, i.e.* $\|\varphi\| = 1_L$ *in each* $\varphi \in T$.

Theorem 1. *(Soundness) Let* L *is linearly ordered ŁΠG$_\triangle$-algebra and* φ *is a formula in* J*, if* $\vdash \varphi$*, then* φ *is* L*-tautology, i.e.* $\|\varphi\|^{\mathrm{L}}_{\mathrm{M}} = 1_{\mathrm{L}}$.

Theorem 2. *(Strong Soundness) Let* T *be a theory,* L *is linearly ordered ŁΠG$_\triangle$-algebra and* φ *is a formula in* J*, if* $T \vdash \varphi$ *(*φ *is provable in* T*), then* $\|\varphi\|^{\mathrm{L}}_{\mathrm{M}} = 1_{\mathrm{L}}$ *for each linearly ordered ŁΠG$_\triangle$-algebra* L *and each* L*-model* M *of* T.

Theorem 3. *(Deduction Theorem) Let* T *be a theory,* φ, ψ *are closed formulas. Then* $(T \cup \{\varphi\}) \vdash \psi$ *iff* $T \vdash \Delta\varphi \to \psi$.

3 Completeness of $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$

Definition 9. *Let T be a theory on $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$.*

(1) T is consistent if there is a formula φ unprovable in T.

(2) T is complete if for each pair φ,ψ of closed formula, $T\vdash(\varphi\to\psi)$ or $T\vdash(\psi\to\varphi)$.

(3) T is Henkin if for each closed formula of the form $(\forall x)\varphi(x)$ unprovable in T there is a constant c in the language of T such that $\varphi(c)$ is unprovable.

Lemma 4. *T is inconsistent iff $T\vdash\overline{0}$.*

Lemma 5. *T is complete iff for each pair φ,ψ of closed formulas if $T\vdash\varphi\vee\psi$, then T proves φ or T proves ψ.*

Proof. Sufficiency: For each pair φ,ψ of closed formulas, being $(\varphi\to\psi)\vee(\psi\to\varphi)$ is theorem in $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$, so $T\vdash(\varphi\to\psi)\vee(\psi\to\varphi)$, thus $T\vdash(\varphi\to\psi)$ or $T\vdash(\psi\to\varphi)$. Thus T is complete. Necessity: assume T is complete and $T\vdash\varphi\vee\psi$, Either $T\vdash\varphi\to\psi$ and then $T\vdash(\varphi\vee\psi)\to\psi$, thus $T\vdash\psi$, or $T\vdash\psi\to\varphi$ and then similarly $T\vdash\varphi$.

Definition 10. *Let T be a theory, the set of all closed formulas over $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$ is denoted by $F^c(\forall\mathcal{UL}^{\triangle}_{h\in(0,1]})$. The definition of relation $\sim_T$ on $F^c(\forall\mathcal{UL}^{\triangle}_{h\in(0,1]})$ is: $\varphi\sim_T\psi$ iff $T\vdash\varphi\to\psi$, $T\vdash\psi\to\varphi$.*

Obviously, $\sim_T$ is equivalent relation on $F^c(\forall\mathcal{UL}^{\triangle}_{h\in(0,1]})$, and holds on &, $\to,\triangle$. So the quotient algebra $[F]_T=F^c(\forall\mathcal{UL}^{\triangle}_{h\in(0,1]})/\sim_T=\{[\varphi]_T|\varphi\in F^c(\forall\mathcal{UL}^{\triangle}_{h\in(0,1]})\}$ of $F^c(\forall\mathcal{UL}^{\triangle}_{h\in(0,1]})$ about $\sim_T$ is Ł$\Pi G_{\triangle}$ algebra, in which, $[\varphi]_T=\{\psi\in F^c(\forall\mathcal{UL}^{\triangle}_{h\in(0,1]})\ |\psi\sim_T\varphi\}$, the partial order $\leq$ on $[F]_T$ is $[\varphi]_T\leq[\psi]_T$ iff $T\vdash\varphi\to\psi$.

Lemma 6. *(1) If T is complete then $[F]_T$ is linearly ordered.*

(2) If T is Henkin then for each formula $\varphi(x)$ with just one free variable x, $[(\forall x)\varphi]_T=\inf_c[\varphi(c)]_T,[(\exists x)\varphi]_T=\sup_c[\varphi(c)]_T$, in which c running over all constants of T.

Proof. (1) is obvious since $[\varphi]_T\leq[\psi]_T$ iff $T\vdash\varphi\to\psi$.

(2) Clearly, $[(\forall x)\varphi(x)]_T\leq\inf_c[\varphi(c)]_T$ for each c, thus $[(\forall x)\varphi(x)]_T\leq\inf_c[\varphi(c)]_T$. To prove that $[(\forall x)\varphi(x)]_T$ is the infimum of all $[\varphi(c)]_T$, assume $[\gamma]_T\leq[\varphi(c)]_T$ for each c, we have to prove $[\gamma]_T\leq[(\forall x)\varphi(x)]_T$(which means that $[(\forall x)\varphi(x)]_T$ is the greatest lower bound of all $[\varphi(c)]_T$). But if $[\gamma]_T\nleq[(\forall x)\varphi(x)]_T$ then $T\nvdash\gamma\to(\forall x)\varphi(x)$, thus $T\nvdash(\forall x)(\gamma\to\varphi(x))$. So by the henkin property, there is a constant c such that $T\nvdash\gamma\to\varphi(c)$, thus $[\gamma]_T\nleq[\varphi(c)]_T$, a contradiction.

Similarly, $[\varphi(c)]_T\leq[(\exists x)\varphi(x)]_T$ for each c. Assume $[\varphi(c)]_T\leq[\gamma]_T$ for each c, we prove $[(\exists x)\varphi]_T\leq[\gamma]_T$. Indeed, if $[(\exists x)\varphi]_T\nleq[\gamma]_T$ then $T\nvdash(\exists x)\varphi(x)\to\gamma$, thus $T\nvdash(\forall x)(\varphi(x)\to\gamma)$ and for some c, $T\nvdash\varphi(c)\to\gamma$, thus $[\varphi(c)]_T\nleq[\gamma]_T$, a contradiction. This completes the proof.

Lemma 7. *For each theory T and each closed formula α, if $T \nvdash \alpha$ then there is a complete Henkin supertheory $\widehat{T}$ of T such that $\widehat{T} \nvdash \alpha$.*

Proof. First observe that if T' is an extension of T, $T' \nvdash \alpha$, and (φ, ψ) is a pair of closed formulas then either $(T' \cup \{\varphi \to \psi\}) \nvdash \alpha$ or $(T' \cup \{\psi \to \varphi\}) \nvdash \alpha$. This is proved easily using the deduction theorem(Theorem 3). Indeed, if $T', \{\varphi \to \psi\} \vdash \alpha$ and $T', \{\psi \to \varphi\} \vdash \alpha$, then $T' \vdash \Delta(\varphi \to \psi) \to \alpha, T' \vdash \Delta(\psi \to \varphi) \to \alpha$, so $T' \vdash \Delta(\varphi \to \psi) \vee \Delta(\psi \to \varphi) \to \alpha$, thus $T' \vdash \alpha$, a contradiction.

Put $T'' = T' \cup \{\varphi \to \psi\}$ in the former case and $T'' = T' \cup \{\psi \to \varphi\}$ in the latter, T'' is the extension of T' deciding (φ, ψ) and keeping α unprovable.

We shall construct $\widehat{T}$ in countably many stages. First extend the language J of T to J' adding new constants $c_0, c_1, c_2, \cdots$. In the construction we have to decide each pair (φ, ψ) of closed J'-formulas and ensure the Henkin property for each closed J'-formula of the form $(\forall x)\chi(x)$. These are countably many tasks and may be enumerated by natural numbers(e.g. in even steps we shall decide all pair (φ, ψ), in odd ones process all formulas $(\forall x)\chi(x)$—or take any other enumeration).

Put $T_0 = T$ and $\alpha_0 = \alpha$, then $T_0 \nvdash \alpha_0$. Assume T_n, α_n have been constructed such that T_n extends T_0, $T_n \vdash \alpha \to \alpha_n, T_n \nvdash \alpha_n$; we construct T_{n+1}, α_{n+1} in such a way that $T_n \vdash \alpha \to \alpha_{n+1}, T_{n+1} \nvdash \alpha_{n+1}$ and T_{n+1} fulfils the n-th task.

Case 1 $n-$th task is deciding (φ, ψ). Let T_{n+1} be extension of T_n deciding (φ, ψ) and keeping α_n unprovable; put $\alpha_{n+1} = \alpha_n$.

Case 2 $n-$th task is processing $(\forall x)\chi(x)$. First let c be one of the new constant not occurring in T_n.

Subcase(a) $T_n \nvdash \alpha_n \vee \chi(c)$, thus $T_n \nvdash (\forall x)\chi(x)$. Put $T_{n+1} = T_n$, $\alpha_{n+1} = \alpha_n \vee \chi(c)$.

Subcase(b) $T_n \vdash \alpha_n \vee \chi(c)$, thus $T_n \vdash \alpha_n \vee \chi(x)$ by the standard argument(in the proof of $\alpha_n \vee \chi(c)$ replace c by a new variable x throughout). Hence $T_n \vdash (\forall x)(\alpha_n \vee \chi(x))$ and using axiom (U17) for the first time, $T_n \vdash \alpha_n \vee (\forall x)\chi(x)$. Thus $T_n \cup \{(\forall x)\chi(x) \to \alpha_n\} \vdash \alpha_n$ so that $T_n \cup \{\alpha_n \to (\forall x)\chi(x)\} \nvdash \alpha_n, T_n \cup \{\alpha_n \to (\forall x)\chi(x)\} \vdash (\forall x)\chi(x)$ does not prove α_n but it does prove $(\forall x)\chi(x)$. Thus put $T_{n+1} = T_n \cup \{\alpha_n \to (\forall x)\chi(x)\}$ and $\alpha_{n+1} = \alpha_n$.

Now let $\widehat{T}$ be the union of all T_n. Then clearly $\widehat{T}$ is complete and $\widehat{T} \vdash \alpha$(since for all n, $\widehat{T} \vdash \alpha$). We show that $\widehat{T}$ is Henkin. Let $\widehat{T} \nvdash (\forall x)\chi(x)$ and let $(\forall x)\chi(x)$ be processed in step n. Then $T_{n+1} \nvdash (\forall x)\chi(x), T_{n+1} \nvdash \alpha_{n+1}$, thus subcase (a) applies and $\widehat{T} \nvdash \alpha_{n+1}$, α_{n+1} being $\alpha_n \vee \chi(c)$. Hence $\widehat{T} \nvdash \chi(c)$. This completes the proof.

Lemma 8. *For each complete Henkin theory T and each closed formula α unprovable in T there is a linearly ordered $\L\Pi G_{\triangle}$-algebra* L *and* L*-model* M *of T such that $\|\alpha\|^{\mathrm{L}}_{\mathrm{M}} < 1_{\mathrm{L}}$.*

Proof. Take M be the set of all constants of the language of T, $m_c = c$ for each such constant. Let L be the lattice of classes of T-equivalent closed formulas, i.e. put $[\varphi]_T = \{\psi | T \vdash \varphi \equiv \psi\}, [\varphi]_T * [\psi]_T = [\varphi \& \psi]_T, [\varphi]_T \Rightarrow [\psi]_T = [\varphi \to \psi]_T$. So L is a linearly ordered ŁΠG$_\triangle$-algebra(since $T \vdash \varphi \to \psi$ or $T \vdash \psi \to \varphi$ for each pair (φ, ψ)).

For each predicate P of arity n, let $r_P(c_1, \cdots, c_n) = [P(c_1, \cdots, c_n)]_T$, this completes the definition of M. It remains to prove $\|\alpha\|^{\mathrm{L}}_{\mathrm{M}} = [\alpha]_T$ for each closed formula φ. Then for each axiom φ of T we have $\|\varphi\|^{\mathrm{L}}_{\mathrm{M}} = [\varphi]_T = [1]_T = 1_{\mathrm{L}}$, but $\|\alpha\|^{\mathrm{L}}_{\mathrm{M}} = [\alpha]_T \neq [1]_T = 1_{\mathrm{L}}$. For atomic closed formula φ the claim follows by definition; the induction step for connectives is obvious. We handle the quantifiers. Let

$(\forall x)\varphi(x), (\exists x)\varphi(x)$ be closed, then by the induction hypothesis,

$\|(\forall x)\varphi(x)\|^{\mathrm{L}}_{\mathrm{M}} = \inf_c \|\varphi(c)\|^{\mathrm{L}}_{\mathrm{M}} = \inf_c [\varphi(c)]_T = [(\forall x)\varphi(x)]_T$

$\|(\exists x)\varphi(x)\|^{\mathrm{L}}_{\mathrm{M}} = \sup_c \|\varphi(c)\|^{\mathrm{L}}_{\mathrm{M}} = \sup_c [\varphi(c)]_T = [(\exists x)\varphi(x)]_T$

Here we use lemma and the fact that in our M, each element c of M is the meaning of a constant (namely itself); this gives $\|(\forall x)\varphi(x)\|^{\mathrm{L}}_{\mathrm{M}} = \inf_c \|\varphi(c)\|^{\mathrm{L}}_{\mathrm{M}}$ and the dual for $\exists$.

Using the above lemmas, we can get the following completeness theorem.

Theorem 4. *(Completeness) For predicate calculus system $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$, T is a theory, φ is a formula, $T \vdash \varphi$ iff for each linearly ordered ŁΠG$_\triangle$-algebra* L *and each safe* L*-model* M *of T, $\|\varphi\|^{\mathrm{L}}_{\mathrm{M}} = 1_{\mathrm{L}}$.*

4 Some Extension Logics of $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$

According to the same method for $h \in (0, 1]$ in propositional system, we can also build up the predicate calculus formal deductive systems $\forall\mathcal{UL}^{\triangle}_{h\in(0,1)}$, $\forall\mathcal{UL}^{\triangle}_{h\in[0.75,1]}$ and $\forall\mathcal{UL}^{\triangle}_{h\in(0,0.75)\cup 1}$ if we fix $h \in (0, 1)$, $h \in [0.75, 1]$ and $h \in (0, 0.75) \cup 1$ respectively in the predicate calculus language J.

Definition 11. *Axioms of the the system $\forall\mathcal{UL}^{\triangle}_{h\in(0,1)}$ are those of $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$ plus*

(ŁΠ) $(\varphi \to (\varphi \& \psi)) \to ((\varphi \to \overline{0}) \vee \psi)$.

Deduction rules of $\forall\mathcal{UL}^{\triangle}_{h\in(0,1)}$ are modus ponens:from $\varphi, \varphi \to \psi$ infer ψ, necessitation: from φ infer $\triangle\varphi$, and generalization: from φ infer $(\forall x)\varphi$.

Definition 12. *Axioms of the the system $\forall\mathcal{UL}^{\triangle}_{h\in[0.75,1]}$ are those of $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$ plus*

(ΠG) $(\varphi \wedge (\varphi \to \overline{0})) \to \overline{0}$.

Deduction rules of $\forall\mathcal{UL}^{\triangle}_{h\in[0.75,1]}$ are modus ponens:from $\varphi, \varphi \to \psi$ infer ψ, necessitation: from φ infer $\triangle\varphi$, and generalization: from φ infer $(\forall x)\varphi$.

Definition 13. *Axioms of the the system $\forall\mathcal{UL}^{\triangle}_{h\in(0,0.75)\cup 1}$ are those of $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$ plus*

(ŁG) $(((\varphi \to \overline{0}) \to \overline{0}) \to \varphi) \vee (\varphi \to (\varphi \& \varphi))$.

Deduction rules of $\forall\mathcal{UL}^{\triangle}_{h\in(0,0.75)\cup 1}$ *are modus ponens:from* $\varphi, \varphi \to \psi$ *infer* ψ*, necessitation: from* φ *infer* $\triangle\varphi$*, and generalization: from* φ *infer* $(\forall x)\varphi$.

Ł$\Pi_{\triangle}$-algebras are ŁΠG$_{\triangle}$-algebras satisfying: $(\varphi \to (\varphi\&\psi)) \to ((\varphi \to \overline{0}) \vee \psi) = 1$. ΠG$_{\triangle}$-algebras are ŁΠG$_{\triangle}$-algebras satisfying $\varphi \wedge (\varphi \to \overline{0}) = 0$. ŁG$_{\triangle}$-algebras are Ł$\Pi$G$_{\triangle}$-algebras satisfying $(((\varphi \to \overline{0}) \to \overline{0}) \to \varphi) \vee (\varphi \to (\varphi\&\varphi)) = 1$.

Remark 2. For each $h \in (0,1), ([0,1], \wedge_h, \Rightarrow_h, \min, \max, 0, 1, \triangle)$ is a linearly ordered Ł$\Pi_{\triangle}$-algebra. For each $h \in [0.75,1], ([0,1], \wedge_h, \Rightarrow_h, \min, \max, 0, 1, \triangle)$ is a linearly ordered ΠG$_{\triangle}$-algebra. For each $h \in (0,0.75) \cup 1, ([0,1], \wedge_h, \Rightarrow_h$ $, \min, \max, 0, 1, \triangle)$ is a linearly ordered ŁG$_{\triangle}$-algebra. So the predicate system $\forall\mathcal{UL}^{\triangle}_{h\in(0,1)}$, $\forall\mathcal{UL}^{\triangle}_{h\in[0.75,1]}$ and $\forall\mathcal{UL}^{\triangle}_{h\in(0,0.75)\cup 1}$ can be considered the axiomatization for 0-level universal AND operator with projection operator according to $h \in (0,1), h \in [0.75,1]$ and $h \in (0,0.75) \cup 1$.

Theorem 5. *(Completeness)For predicate calculus system* $\forall\mathcal{UL}^{\triangle}_{h\in(0,1)}$*,* T *is a theory,* φ *is a formula,* $T \vdash \varphi$ *iff for each linearly ordered Ł$\Pi_{\triangle}$-algebra* L *and each safe* L*-model* M *of* T*,* $\|\varphi\|^{\mathrm{L}}_{\mathrm{M}} = 1_{\mathrm{L}}$.

Theorem 6. *(Completeness) For predicate calculus system* $\forall\mathcal{UL}^{\triangle}_{h\in[0.75,1]}$*,* T *is a theory,* φ *is a formula,* $T \vdash \varphi$ *iff for each linearly ordered ΠG$_{\triangle}$-algebra* L *and each safe* L*-model* M *of* T*,* $\|\varphi\|^{\mathrm{L}}_{\mathrm{M}} = 1_{\mathrm{L}}$.

Theorem 7. *(Completeness) For predicate calculus system* $\forall\mathcal{UL}^{\triangle}_{h\in(0,0.75)\cup 1}$*,* T *is a theory,* φ *is a formula,* $T \vdash \varphi$ *iff for each linearly ordered ŁG$_{\triangle}$-algebra* L *and each safe* L*-model* M *of* T*,* $\|\varphi\|^{\mathrm{L}}_{\mathrm{M}} = 1_{\mathrm{L}}$.

5 Conclusion

In this paper a predicate calculus formal deductive system $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$ according to the propositional system $\mathcal{UL}^{\triangle}_{h\in(0,1]}$ for 0-level universal AND operator is introduced. We prove the system $\mathcal{UL}^{\triangle}_{h\in(0,1]}$ is complete. The completeness of some extension logics of $\forall\mathcal{UL}^{\triangle}_{h\in(0,1]}$ are also given. The following work for us is to study the axiomatization for universal logic in other cases.

Acknowledgements. This work is supported by the Special Scientific Research Plan Project of Shaanxi Provincial Education Department(No. 2010JK567).

References

1. Hajek, P.: Metamathematics of fuzzy logic. Kluwer Academic Publishers, Dordrecht (1998)
2. Esteva, F., Godo, L.: Monoidal t-normbased logic: Towards a logic for left-continous t-norms. Fuzzy Sets and Systems 124, 271–288 (2001)
3. Cignoli, R., Esteva, F., Godo, L.: Basic fuzzy logic is the logic of continuous t-norms and their residua. Soft computing 4, 106–112 (2000)
4. Hohle, U.: Commutative residuated l-monoids. In: Hohle, U., Klement, E.P. (eds.) Non-classical logics and their applications to fuzzy subsets. Kluwer Academic Publishers, Dordrecht (1995)
5. Esteva, F., Godo, L., et al.: Residuated fuzzy logics with an involutive negation. Archive for Mathmatical Logic 39, 103–124 (2000)
6. Klement, E.P., Mesiar, R., Pap, E.: Triangular norms. Kluwer Academic Publishers, Dordrecht (2000)
7. Pei, D.W., Wang, G.J.: The completeness and applications of the formal system L*. Science in China (Series F) 45, 40–50 (2002)
8. Wang, S.M., Wang, B.S., Pei, D.W.: A fuzzy logic for an ordinal sum t-norm. Fuzzy Sets and Systems 149, 297–307 (2005)
9. Wang, G.J.: Non-classical Mathematical logic and approximate reasoning. Science Press, Beijing (2000) (in Chinese)
10. Pei, D.W.: First-order formal system K^* and its completeness. Chinese Annals of Mathematics, Series A 23, 675–684 (2002)
11. Wu, H.B.: Competeness of $BL^*_\triangle$ system. Journal of Jishou University (Natural Science Edition) 30(2), 1–5 (2009)
12. He, H.C., et al.: Universal logic principle. Science Press, Beijing (2001) (in Chinese)
13. Ma, Y.C., He, H.C.: The fuzzy reasoning rules based on universal logic. In: 2005 IEEE International Conference on GrC, pp. 561–564 (2005)
14. Ma, Y.C., He, H.C.: A propositional calculus formal deductive system $\mathcal{UL}_{h\in(0,1]}$ of universal logic. In: Proceedings of 2005 ICMLC, pp. 2716–2721 (2005)
15. Ma, T.C., Li, Q.Y.: A propositional deductive system Of universal logic with projection operator. In: Proceedings of 2006 ISDA, pp. 993–998 (2006)
16. Ma, Y.C., He, H.C.: The axiomatization for 0-level universal logic. In: Yeung, D.S., Liu, Z.-Q., Wang, X.-Z., Yan, H. (eds.) ICMLC 2005. LNCS (LNAI), vol. 3930, pp. 367–376. Springer, Heidelberg (2006)
17. Ma, Y.C., He, H.C.: Axiomatization for 1-level universal and operator. The Journal of China Universities of Posts and Telecommunications 15, 125–129 (2008)
18. Ma, Y.C., Zhang, M.: A predicate formal system of universal logic with projectionoperator. Accepted in Lecture Notes in Computer Sciences/Artificial Intelligence (2010)

The Degree of the Quasi-similarity and the Pseudo-metric between Predicate Formulae

Xiao-yan Qin[1,2], Yi Liu[2,3], and Yang Xu[2]

[1] College of Mathematics and Computer Science, Shanxi Normal University, Linfen 041004, Shanxi, China
e-mail: lisaqin1981@126.com

[2] Intelligent Control Development Center, Southwest Jiaotong University, Chengdu 610031, Sichuan, China
e-mail: {lisaqin1981,liuyiyl}@126.com, xuyang@swjtu.edu.cn

[3] Key Laboratory of Numerical Simulation in the Sichuan Provincial College, Neijiang Normal University, Neijiang 641000, Sichuan, China
e-mail: liuyiyl@126.com

Abstract. The theory of the quasi-truth degrees of predicate formulae is a preliminary test of quantitative predicate logic. Corresponding to quantitative propositional logic, we know that it's just a beginning. In this paper, we continue the research of the measurement in predicate logic.We propose the concept of the degree of the quasi-similarity and logically quasi-equivalent relation between formulae, and study some important reasonable results. Moreover, the pseudo-metric ρ on the set $\mathcal{F}$ of formulae is naturally induced, and we prove that the operators $\vee$, $\wedge$ and $\rightarrow$ are all continuous on the logic metric space $(\mathcal{F}, \rho)$. The paper further riches the theory of quantitative predicate logic, and provides a basic framework for the approximate reasoning in predicate logic.

Keywords: Quantitative predicate logic, approximate reasoning, quasi-similarity degree, pseudo-metric.

1 Introduction

It is well known that the distinguished features of the subject of the mathematical logic are symbolization and formalization that endow this subject with a special style completely different from the style of computation mathematics.In fact,the former lays stresses on formal deduction, while the latter attaches importance to numerical calculation;the former puts emphasis on rigid proof,while the latter tolerates approximate computation.It seems that there exists an invisible separation wall between them.

In 2009, Wang [1] proposed quantitative logic which combined the mathematical logic and the computation mathematics to provide a graded approach

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 273–283.
springerlink.com

to mathematical logic.Quantitative logic is composed of quantitative propositional logic and quantitative predicate logic,and the elementary theories of quantitative propositional logic include five parts: the degree of the truth of formulae,the degree of the similarity between formulae,logic metric space $(F(S), \rho)$,theory of the approximate reasoning in $F(S)$,the degree of the consistency of theories.

There are many researches on quantitative propositional logic up to now([1-6]), while it's a close blanket in the case of quantitative predicate logic. It's because of the complexity of the interpretations of the first-order language that we find it hard to define the truth degrees of first-order formulae,which is the base theory of quantitative logic. Then, can we only consider the interpretations with finite domain in some approximate sense? It's a pity that the answer is "No". Hilbert ever gave the following formula A^*:

$$(\exists x)\, P(x,x) \vee (\exists x)(\exists y)(\exists z)\,(P(x,y) \wedge P(y,z) \wedge \neg P(x,z)) \vee (\exists x)(\forall y)\, \neg P(y,x)\,.$$

Here A^* is true w.r.t. any interpretation with finite domain, but there exists an interpretation with a infinite domain w.r.t. which A^* is false([7]). Even so, if a formula is true w.r.t. any interpretation with a finite domain, it's certainly a good formula to some extent. In this way, the authors proposed the degree of the quasi-truth of the first-order formula to measure approximately its truth degree in [8].

However, it's only a beginning of quantitative predicate logic. In this paper, we will furthermore study the elementary theories of quantitative predicate logic in this viewpoint. Concretely, we recall some basic theories of the quasi-truth degrees of formulae as a preliminary part in section 2. In section 3, we propose the definition of the degree of the quasi-similarity between formulae, and study some important properties. Moreover, we propose the concept of logically quasi-equivalent relation between formulae to discuss some further algebraic properties. In section 4, a kind of pseudo-metric on the set $\mathcal{F}$ of formulae is given, and we prove that the operations such as $\neg$, $\rightarrow$, $\vee$, and $\wedge$ are all continuous on the metric space $(\mathcal{F}, \rho)$. Based on the pseudo-metric on the set $\mathcal{F}$ of formulae, we can make further studies on the theory of approximate reasoning. We leave this part of development to subsequent work.Section 5 is the conclusion.

2 Preliminaries

2.1 Basic Concepts and Symbols in the Present Paper

A first order language $\mathcal{L}$ has the following as its alphabet of symbols:(i) variables $x_1, x_2, \cdots$; (ii) some individual constants $c_1, c_2, \cdots$;(iii) some predicate letters P,Q,R,$\cdots$;(iv) some function letters $f, g, h, \cdots$;(v) the punctuation symbols "(",")" and ",";(vi) the connectives $\neg$ and $\rightarrow$;(vii) the universal quantifier $\forall$; (viii) the existential quantifier $\exists$, which is an abbreviation of

$\neg\forall\neg$(see, e.g. [9]). In the present paper, we agree on the assumption that there is no function symbols in the alphabet, which is adopted in, for example, Refs. [10,11] etc. Then $x_1, x_2, \cdots$ and $c_1, c_2, \cdots$ are called terms. An atomic formula has the form $P(t_1, \cdots, t_k)$ where P is a predicate letter of arity k and $t_1, \cdots, t_k$ are terms. A well-formed formula(briefly wff or first-order formula, or simply formula) of $\mathcal{L}$ is defined by:(i) Each atomic formula of $\mathcal{L}$ is a formula. (ii) If A and B are formulae of $\mathcal{L}$, then so are $\neg A, A \to B$ and $(\forall x_i)A$. (iii) The set of formulae of $\mathcal{L}$ is generated as in (i) and (ii). The set of all formulae of $\mathcal{L}$ will be denoted by $\mathcal{F}$.

An interpretation $\mathbf{I} = (I, (r_P)_P, (m_c)_c)$ (or, 0,1-structure called by Há́jek in [11]) is a triplet where I is a nonempty set, called the domain of $\mathbf{I}$, for each predicate letter P of arity k, r_P is a relation of arity k on I, i.e., $r_P \subseteq I^k$, and for each individual constant c, there exists a unique corresponding element m_c in I. A valuation v of $\mathcal{L}$ in $\mathbf{I}$ is a mapping from the set $\mathcal{T}$ of terms to I satisfying $v(c) = m_c$ and $v(x_n) \in I(n = 1, 2, \cdots)$. A valuation v in $\mathbf{I}$ is said to satisfy a formula $A \in \mathcal{F}$ if it can be shown inductively to do so under the following four conditions: (i) v satisfies the atomic formula $P(t_1, t_2, \cdots, t_k)$ if $r_P(t_1, t_2, \cdots, t_k)$ is true in $\mathbf{I}$, i.e., $(v(t_1), v(t_2), \cdots, v(t_k)) \in r_P$. (ii) v satisfies $\neg B$ if v does not satisfy B. (iii) v satisfies $B \to C$ if v satisfies $\neg B$ or v satisfies C. (iv) v satisfies $(\forall x_i)B$ if every valuation $v^{'}$ which is i-equivalent to v satisfies B, where $v^{'}$ is i-equivalent to v iff $v^{'}(x_n) = v(x_n)$ whenever $n \neq i$. The set of all valuations in $\mathbf{I}$ is denoted by $\Omega_{\mathbf{I}}$. Assume that $A \in \mathcal{F}$ and $v \in \Omega_{\mathbf{I}}$, then we use $\|A\|_{\mathbf{I},v} = 1$ to denote that v satisfies A and use $\|A\|_{\mathbf{I},v} = 0$ to denote that v doesn't satisfy A. If $\|A\|_{\mathbf{I},v} = 1$ holds for all $v \in \Omega_{\mathbf{I}}$, then we say that A is true in $\mathbf{I}$, denoted by $\|A\|_{\mathbf{I}} = 1$. If A is true in every interpretation $\mathbf{I}$, then we say that A is logically valid.

2.2 The Degree of the Quasi-truth of Formulae

Definition 1. *([12]) Suppose that $(X_n, \mathcal{A}_n, \mu_n)$ is a probabilistic measure space where μ_n is a probability measure on X_n, and $\mathcal{A}_n$ is the family consisting of all μ_n-measurable subsets of $X_n (n = 1, 2, \cdots)$. Assume that $X = \prod_{n=1}^{\infty} X_n$, then $\prod_{n=1}^{\infty} \mathcal{A}_n$ generates on X a σ-algebra $\mathcal{A}$, and there exists on X a unique measure μ such that (i)$\mathcal{A}$ is the family of all μ-measurable subsets of X; (ii)$E \times \prod_{n=m+1}^{\infty} X_n$ is μ-measurable and*

$$\mu(E \times \prod_{n=m+1}^{\infty} X_n) = (\mu_1 \times \mu_2 \times \cdots \times \mu_m)(E), m = 1, 2, \cdots \qquad (1)$$

whenever E is a measurable subset of $\prod_{n=1}^{m} X_n$. μ is called the infinite product of $\mu_1, \mu_2, \cdots$. The probability measure space $(X, \mathcal{A}, \mu)$ will often be simplified as X.

Suppose that $\mathbf{I} \in \mathcal{I}_f$, $X_n = I$, μ_n is the evenly distributed probability measure on X_n. Let $X_{\mathbf{I}} = \prod_{n=1}^{\infty} X_n$ and $\mu_{\mathbf{I}}$ be the infinite product of $\mu_1, \mu_2, \cdots$. Suppose that $\upsilon \in \Omega_{\mathbf{I}}$, then υ is decided by its restriction $\upsilon|S$ on the set $S = \{x_1, x_2, \cdots\}$ of variables because $(m_c)_c$ in $\mathbf{I}$ is fixed. Assume that $\upsilon(x_k) = \upsilon_k (k = 1, 2, \cdots)$, then $\mathbf{v} = (\upsilon_1, \upsilon_2, \cdots) \in X_{\mathbf{I}}$. Conversely, assume that $\mathbf{v} = (\upsilon_1, \upsilon_2, \cdots) \in X_{\mathbf{I}}$, then there exists a unique $\upsilon \in \Omega_{\mathbf{I}}$ such that $\upsilon(x_k) = \upsilon_k (k = 1, 2, \cdots)$. Hence $\varphi : \Omega_{\mathbf{I}} \to X_{\mathbf{I}}$ is a bijection where $\varphi(\upsilon) = \mathbf{v}$.

Definition 2. *([8]) Suppose that* $\mathbf{I} \in \mathcal{I}_f$ *and* $A \in \mathcal{F}$. *Define* $[A]_{\mathbf{I}}$ *and* $\tau_{\mathbf{I}}(A)$ *as follows respectively*

$$[A]_{\mathbf{I}} = \{\mathbf{v} \in X_{\mathbf{I}} | \upsilon \in \Omega_{\mathbf{I}}, \|A\|_{\mathbf{I},\upsilon} = 1\}, \tau_{\mathbf{I}}(A) = \mu_{\mathbf{I}}([A]_{\mathbf{I}}), \tag{2}$$

then $\tau_{\mathbf{I}}(A)$ *is called the relative truth degree of* A *in* $\mathbf{I}$.

Definition 3. *([8]) Suppose that* $A \in \mathcal{F}$. *Define* $\tau(A)$ *as follows:*

$$\tau(A) = \frac{\sup\{\tau_{\mathbf{I}}(A) | \mathbf{I} \in \mathcal{I}_f\} + \inf\{\tau_{\mathbf{I}}(A) | \mathbf{I} \in \mathcal{I}_f\}}{2} \tag{3}$$

then $\tau(A)$ *is called the degree of the quasi-truth of* A.

Remark 1. The above degree is just an average truth degree of A, and we can obtain other measurement through dealing with all the relative truth degrees of A in other ways, e.g. weighting method. We leave this work in the consequent paper.

Proposition 1. *Suppose that* $A, B \in \mathcal{F}$.
(i) $0 \leq \tau(A) \leq 1$;
(ii) $\tau(A) + \tau(\neg A) = 1$;
(iii) $\tau(A \wedge B) \leq \tau(A) \wedge \tau(B)$;
(iv) $\tau(A) \vee \tau(B) \leq \tau(A \vee B)$.

Theorem 1. *([8]) Suppose that* $A, B \in \mathcal{F}$, B *is a prenex normal form, and* A *is logically equivalent to* B. *If* B *contains no quantifiers, or the quantifiers in* B *are all universal or all existential, then*
(i) A *is logically efficient if and only if* $\tau(A) = 1$;
(ii) A *is a contradiction if and only if* $\tau(A) = 0$.

Remark 2. It is easy to verify that if A is logically valid then $\tau(A) = 1$, and if A is a contradiction, then $\tau(A) = 0$, where no extra conditions are required.

Theorem 2. *([8]) Suppose that* $A, B, C \in \mathcal{F}$.
(i) If $\tau(A) \geq \alpha, \tau(A \to B) \geq \beta$, *then* $\tau(B) \geq \alpha + \beta - 1$;
(ii) If $\tau(A \to B) \geq \alpha, \tau(B \to C) \geq \beta$, *then* $\tau(A \to C) \geq \alpha + \beta - 1$.

Theorem 3. *([8]) Suppose that* $A \in \mathcal{F}$.
(i) $\tau((\forall x)A) \leq \tau(A)$;
(ii) $\tau(A) = 1$ *if and only if* $\tau((\forall x)A) = 1$.

Remark 3. From Theorem 2 and Theorem 3, we know that MP rule,HS rule and the Generalization rule all conserve the property that the degree of the quasi-truth equals to 1.

In [13], Wang proved that the following theorem held, and it shows us that it's feasible to reason approximately on the set $\mathcal{F}$ of all the formulae based on the degrees of the quasi-truth of formulae.

Theorem 4. *The set of all the degrees of the quasi-truth of formulae is dense in [0,1].*

3 The Degree of the Quasi-similarity between Formulae

3.1 The Relative Degree of the Similarity between Formulae

Definition 4. *Suppose that* $\mathbf{I} \in \mathcal{I}_f$, *and* $A, B \in \mathcal{F}$. *Define* $\xi_{\mathbf{I}}(A,B)$ *as follows:*

$$\xi_{\mathbf{I}}(A,B) = \tau_{\mathbf{I}}((A \to B) \wedge (B \to A)). \tag{4}$$

then $\xi_{\mathbf{I}}(A,B)$ *is called the relative degree of the similarity between* A *and* B *in* $\mathbf{I}$.

Proposition 2. *Suppose that* $\mathbf{I} \in \mathcal{I}_f$, *and* $A, B \in \mathcal{F}$.
(i) $0 \le \xi_{\mathbf{I}}(A,B) \le 1$;
(ii) $\xi_{\mathbf{I}}(A,B) = \mu_{\mathbf{I}}(\{\mathbf{v} \in X_{\mathbf{I}} | v \in \Omega_{\mathbf{I}}, \| A \|_{\mathbf{I},v} = \| B \|_{\mathbf{I},v}\})$;
(iii) $\xi_{\mathbf{I}}(A,B) + \xi_{\mathbf{I}}(A,\neg B) = 1$.

Proposition 3. *Suppose that* $\mathbf{I} \in \mathcal{I}_f$, *and* $A, B, C, D \in \mathcal{F}$.
(i) $\xi_{\mathbf{I}}(A,B) + \xi_{\mathbf{I}}(B,C) \le 1 + \xi_{\mathbf{I}}(A,C)$;
(ii) $\xi_{\mathbf{I}}(A,C) + \xi_{\mathbf{I}}(B,D) \le 1 + \xi_{\mathbf{I}}(A \to B, C \to D)$;
(iii) $\xi_{\mathbf{I}}(A,(\forall x)A) = \tau_{\mathbf{I}}(A \to (\forall x)A)$.

Proof. (i) According to the fact that $\mu(Y \cup Z) = \mu(Y) + \mu(Z) - \mu(Y \cap Z)$, we can prove (i) easily.

(ii) Let

$$G_1 = \{\mathbf{v} \in X_{\mathbf{I}} | v \in \Omega_{\mathbf{I}}, \|A\|_{\mathbf{I},v} = \|C\|_{\mathbf{I},v}\},$$

$$G_2 = \{\mathbf{v} \in X_{\mathbf{I}} | v \in \Omega_{\mathbf{I}}, \|B\|_{\mathbf{I},v} = \|D\|_{\mathbf{I},v}\},$$

$$G_3 = \{\mathbf{v} \in X_{\mathbf{I}} | v \in \Omega_{\mathbf{I}}, \|A \to B\|_{\mathbf{I},v} = \|C \to D\|_{\mathbf{I},v}\},$$

then it's clear that $G_1 \cap G_2 \subseteq G_3$, and $G_1 \cup G_2 \subseteq X_{\mathbf{I}}$. So we have

$$\begin{aligned}
\xi_{\mathbf{I}}(A,C) + \xi_{\mathbf{I}}(B,D) &= \mu_{\mathbf{I}}(G_1) + \mu_{\mathbf{I}}(G_2) \\
&= \mu_{\mathbf{I}}(G_1 \cup G_2) + \mu_{\mathbf{I}}(G_1 \cap G_2) \\
&\le \mu_{\mathbf{I}}(X) + \mu_{\mathbf{I}}(G_3) \\
&= 1 + \xi_{\mathbf{I}}(A \to B, C \to D).
\end{aligned}$$

(iii) Because the formula $(\forall x)A \to A$ is logically efficient, we can easily prove that $\tau_{\mathbf{I}}((A \to (\forall x)A) \wedge ((\forall x)A \to A)) = \tau_{\mathbf{I}}(A \to (\forall x)A)$ holds for every finite interpretation $\mathbf{I}$.Thus we have $\xi_{\mathbf{I}}(A, (\forall x)A) = \tau_{\mathbf{I}}(A \to (\forall x)A)$ from (4).

Corollary 1. *For any $A, B \in \mathcal{F}$ and $\mathbf{I} \in \mathcal{I}_f$, if $\xi_{\mathbf{I}}(A, B) = \xi_{\mathbf{I}}(B, C) = 1$, then $\xi_{\mathbf{I}}(A, C) = 1$.*

3.2 The Degree of the Quasi-similarity between Formulae

Based on the the relative degree of the similarity between formulae,we can define the degree of the quasi-similarity between formulae as follows:

Definition 5. *Suppose that $A, B \in \mathcal{F}$,Define $\xi(A, B)$ as follows:*

$$\xi(A, B) = \frac{\sup\{\xi_{\mathbf{I}}(A, B)|\mathbf{I} \in \mathcal{I}_f\} + \inf\{\xi_{\mathbf{I}}(A, B)|\mathbf{I} \in \mathcal{I}_f\}}{2}. \tag{5}$$

then $\xi(A, B)$ is called the degree of the quasi-similarity between A and B.

Example 1. Calculate the quasi-similarity degree between $P(x)$ and $(\forall x)P(x)$.

Solution. For any interpretation $\mathbf{I} = (I, (r_P)_P, (m_c)_c)$ in $\mathcal{I}_f$, it follows from Proposition 3(iii) and (2) that

$$\begin{aligned} \xi_{\mathbf{I}}(P(x), (\forall x)P(x)) &= \tau_{\mathbf{I}_0}(P(x) \to (\forall x)P(x)) \\ &= \mu_{\mathbf{I}}\{\mathbf{v} \in X_{\mathbf{I}} | v \in \Omega_{\mathbf{I}}, \|P(x) \to (\forall x)P(x)\|_{\mathbf{I},v} = 1\} \\ &= 1 - \mu_{\mathbf{I}}\{\mathbf{v} \in X_{\mathbf{I}} | v \in \Omega_{\mathbf{I}}, \|P(x)\|_{\mathbf{I},v} = 1, \|(\forall x)P(x)\|_{\mathbf{I},v} = 0\}. \end{aligned}$$

So we have that

$$\xi_{\mathbf{I}}(P(x), (\forall x)P(x)) = \begin{cases} 1 & , \ \text{if } r_P = I \text{ or } \emptyset, \\ 1 - \frac{|r_P|}{|I|} & , \ \text{otherwise.} \end{cases}$$

Thus it's clear that

$$\sup\{\xi_{\mathbf{I}}(P(x), (\forall x)P(x))|\mathbf{I} \in \mathcal{I}_f\} = 1.$$

On the other hand, for any positive integer n, construct an interpretation $\mathbf{I}_n = (I_n, (r_{P,n})_P, (m_c)_c)$ such that $|I_n| = n$ and $|r_{P,n}| = n - 1$. Then for the sequence $\{\mathbf{I}_n\}_{n=1}^{\infty}$ of interpretations, $\lim\limits_{n\to\infty} \frac{|r_{P,n}|}{|I|} = 1$. So we have that $\inf\{\xi_{\mathbf{I}}(P(x), (\forall x)P(x))|\mathbf{I} \in \mathcal{I}_f\} = 0$. It follows from (5) that $\xi(P(x), (\forall x)P(x)) = \frac{1}{2}$.

Proposition 4. *Suppose that $A, B \in \mathcal{F}$.*
(i) $0 \leq \xi(A, B) \leq 1$;
(ii) $\xi(A, B) = 1$ if and only if $\forall \mathbf{I} \in \mathcal{I}_f, \forall v \in \Omega_{\mathbf{I}}, \|A\|_{\mathbf{I},v} = \|B\|_{\mathbf{I},v}$;
(iii) $\xi(A, B) = \tau((A \to B) \wedge (B \to A))$;
(iv) $\xi(A, B) \leq \tau(A \to B) \wedge \tau(B \to A)$.

Theorem 5. *Suppose that* $A, B \in \mathcal{F}$.
(i) $\xi(A, A) = 1; \xi(A, \neg A) = 0$;
(ii) $\xi(A, B) = \xi(B, A)$;
(iii) $\xi(A, B) + \xi(A, \neg B) = 1$.

Lemma 1. *For any* $A, B \in \mathcal{F}$, *there exists a finite interpretation* $\mathbf{I} \in \mathcal{I}_f$, *such that* $\xi_{\mathbf{I}}(A, B) \in \{0, 1\}$.

Proof. Let $\mathbf{I} = (I, (r_P)_P, (m_c)_c) \in \mathcal{I}_f$, where $r_P = I^k$ for any predicate letter P of arity k. Then either $\|(A \to B) \wedge (B \to A)\|_{\mathbf{I},v} = 1$ holds for every valuation v in $\Omega_{\mathbf{I}}$, or $\|(A \to B) \wedge (B \to A)\|_{\mathbf{I},v} = 0$ holds for every valuation v in $\Omega_{\mathbf{I}}$, and hence it follows from (2) and (4) that either $\xi_{\mathbf{I}}(A, B) = 1$ or $\xi_{\mathbf{I}}(A, B) = 0$, i.e., $\xi_{\mathbf{I}}(A, B) \in \{0, 1\}$.

Theorem 6. *Suppose that* $A, B, C, D \in \mathcal{F}$.
(i) $\xi(A, B) + \xi(B, C) \leq 1 + \xi(A, C)$;
(ii) $\xi(A, C) + \xi(B, D) \leq 1 + \xi(A \to B, C \to D)$;
(iii) $\xi(A, (\forall x)A) = \tau(A \to (\forall x)A)$.

Proof. Construct an interpretation $\mathbf{I}_0$ in $\mathcal{I}_f$ such that $r_P = {I_0}^k$ in $\mathbf{I}_0$ holds for every predicate symbol P of arity k.

(i) It's clear that we have $\xi_{\mathbf{I}_0}(A, B) \in \{0, 1\}$, and $\xi_{\mathbf{I}_0}(B, C) \in \{0, 1\}$.

① If $\xi_{\mathbf{I}_0}(A, B) = 1$, and $\xi_{\mathbf{I}_0}(B, C) = 1$, then $\xi_{\mathbf{I}_0}(A, C) = 1$ holds according to Corollary 1. It follows from (5) that

$$\inf\{\xi_{\mathbf{I}}(A, B) | \mathbf{I} \in \mathcal{I}_f\} = 2\xi(A, B) - 1, \tag{6}$$

$$\inf\{\xi_{\mathbf{I}}(B, C) | \mathbf{I} \in \mathcal{I}_f\} = 2\xi(B, C) - 1. \tag{7}$$

From (6),(7) and Proposition 3(i), we have

$$\begin{aligned} \inf\{\xi_{\mathbf{I}}(A, C) | \mathbf{I} \in \mathcal{I}_f\} &\geq \inf\{\xi_{\mathbf{I}}(A, B) + \xi_{\mathbf{I}}(B, C) - 1 | \mathbf{I} \in \mathcal{I}_f\} \\ &\geq \inf\{\xi_{\mathbf{I}}(A, B) | \mathbf{I} \in \mathcal{I}_f\} + \inf\{\xi_{\mathbf{I}}(B, C) | \mathbf{I} \in \mathcal{I}_f\} - 1 \\ &= 2\xi(A, B) + 2\xi(B, C) - 3. \end{aligned}$$

Therefore we have

$$\begin{aligned} \xi(A, C) &= \tfrac{\sup\{\xi_{\mathbf{I}}(A, C) | \mathbf{I} \in \mathcal{I}_f\} + \inf\{\xi_{\mathbf{I}}(A, C) | \mathbf{I} \in \mathcal{I}_f\}}{2} \\ &= \tfrac{1 + \inf\{\xi_{\mathbf{I}}(A, C) | \mathbf{I} \in \mathcal{I}_f\}}{2} \\ &\geq \tfrac{1 + 2\xi(A, B) + 2\xi(B, C) - 3}{2} \\ &= \xi(A, B) + \xi(B, C) - 1. \end{aligned}$$

② If $\xi_{\mathbf{I}_0}(A, B) = 1$ and $\xi_{\mathbf{I}_0}(B, C) = 0$, it's clear that $\xi_{\mathbf{I}_0}(A, C) = 0$ according to Proposition 2(ii). It follows from (5) that

$$\inf\{\xi_{\mathbf{I}}(A, B) | \mathbf{I} \in \mathcal{I}_f\} = 2\xi(A, B) - 1, \tag{8}$$

$$\sup\{\xi_{\mathbf{I}}(B, C) | \mathbf{I} \in \mathcal{I}_f\} = 2\xi(B, C) \tag{9}$$

From (8),(9) and Proposition 3(i), we have

$$\begin{aligned}\sup\{\xi_{\mathbf{I}}(A,C)|\mathbf{I}\in\mathcal{I}_f\} &\geq \sup\{\xi_{\mathbf{I}}(A,B)+\xi_{\mathbf{I}}(B,C)-1|\mathbf{I}\in\mathcal{I}_f\}\\ &\geq \sup\{2\xi(A,B)-1+\xi_{\mathbf{I}}(B,C)-1|\mathbf{I}\in\mathcal{I}_f\}\\ &= 2\xi(A,B)-2+\sup\{\xi_{\mathbf{I}}(B,C)|\mathbf{I}\in\mathcal{I}_f\}\\ &= 2\xi(A,B)+2\xi(B,C)-2\end{aligned}$$

Therefore we have

$$\begin{aligned}\xi(A,C) &= \tfrac{\sup\{\xi_{\mathbf{I}}(A,C)|\mathbf{I}\in\mathcal{I}_f\}+\inf\{\xi_{\mathbf{I}}(A,C)|\mathbf{I}\in\mathcal{I}_f\}}{2}\\ &= \tfrac{\sup\{\xi_{\mathbf{I}}(A,C)|\mathbf{I}\in\mathcal{I}_f\}+0}{2}\\ &\geq \tfrac{2\xi(A,B)+2\xi(B,C)-2+0}{2}\\ &= \xi(A,B)+\xi(B,C)-1.\end{aligned}$$

③ If $\xi_{\mathbf{I}_0}(A,B)=0$,and $\xi_{\mathbf{I}_0}(B,C)=1$, then the conclusion can be proved similarly to ②.

④ If $\xi_{\mathbf{I}_0}(A,B)=\xi_{\mathbf{I}_0}(B,C)=0$, then $\xi(A,B)\leq\frac{1}{2}$ and $\xi(B,C)\leq\frac{1}{2}$ hold according to (5). So we have $\xi(A,C)\geq 0=\frac{1}{2}+\frac{1}{2}-1\geq\xi(A,B)+\xi(B,C)-1$.

(ii) Similar to (i), we have $\xi_{\mathbf{I}_0}(A,C)\in\{0,1\},\xi_{\mathbf{I}_0}(B,D)\in\{0,1\}$ and $\xi_{\mathbf{I}_0}(A\to B,C\to D)\in\{0,1\}$.

① If $\xi_{\mathbf{I}_0}(A,C)=1$ and $\xi_{\mathbf{I}_0}(B,D)=1$, then $\xi_{\mathbf{I}_0}(A\to B,C\to D)=1$, then it follows from (5) and Proposition 3(ii) that

$$\begin{aligned}\xi(A,C)+\xi(B,D) &= \tfrac{1+\inf\{\xi_{\mathbf{I}}(A,C)|\mathbf{I}\in\mathcal{I}_f\}}{2}+\tfrac{1+\inf\{\xi_{\mathbf{I}}(B,D)|\mathbf{I}\in\mathcal{I}_f\}}{2}\\ &= 1+\tfrac{\inf\{\xi_{\mathbf{I}}(A,C)|\mathbf{I}\in\mathcal{I}_f\}+\inf\{\xi_{\mathbf{I}}(B,D)|\mathbf{I}\in\mathcal{I}_f\}}{2}\\ &\leq 1+\tfrac{\inf\{\xi_{\mathbf{I}}(A,C)+\xi_{\mathbf{I}}(B,D)|\mathbf{I}\in\mathcal{I}_f\}}{2}\\ &\leq 1+\tfrac{\inf\{1+\xi_{\mathbf{I}}(A\to B,C\to D)|\mathbf{I}\in\mathcal{I}_f\}}{2}\\ &= 1+\tfrac{1+\inf\{\xi_{\mathbf{I}}(A\to B,C\to D)|\mathbf{I}\in\mathcal{I}_f\}}{2}\\ &= 1+\xi(A\to B,C\to D).\end{aligned}$$

② If $\xi_{\mathbf{I}_0}(A,C)=1$, $\xi_{\mathbf{I}_0}(B,D)=0$, and $\xi_{\mathbf{I}_0}(A\to B,C\to D)=0$, then it follows from (5) and Proposition 3(ii) that

$$\begin{aligned}\xi(A,C)+\xi(B,D) &= \tfrac{1+\inf\{\xi_{\mathbf{I}}(A,C)|\mathbf{I}\in\mathcal{I}_f\}}{2}+\tfrac{\sup\{\xi_{\mathbf{I}}(B,D)|\mathbf{I}\in\mathcal{I}_f\}+0}{2}\\ &= \tfrac{1}{2}+\tfrac{\inf\{\xi_{\mathbf{I}}(A,C)|\mathbf{I}\in\mathcal{I}_f\}+\sup\{\xi_{\mathbf{I}}(B,D)|\mathbf{I}\in\mathcal{I}_f\}}{2}\\ &\leq \tfrac{1}{2}+\tfrac{\sup\{\xi_{\mathbf{I}}(B,D)+\inf\{\xi_{\mathbf{I}}(A,C)|\mathbf{I}\in\mathcal{I}_f\}|\mathbf{I}\in\mathcal{I}_f\}}{2}\\ &\leq \tfrac{1}{2}+\tfrac{\sup\{\xi_{\mathbf{I}}(A,C)+\xi_{\mathbf{I}}(B,D)|\mathbf{I}\in\mathcal{I}_f\}}{2}\\ &\leq \tfrac{1}{2}+\tfrac{\sup\{1+\xi_{\mathbf{I}}(A\to B,C\to D)\}}{2}\\ &= 1+\tfrac{\sup\{\xi_{\mathbf{I}}(A\to B,C\to D)|\mathbf{I}\in\mathcal{I}_f\}+0}{2}\\ &= 1+\xi(A\to B,C\to D).\end{aligned}$$

③ If $\xi_{\mathbf{I}_0}(A,C)=0$, $\xi_{\mathbf{I}_0}(B,D)=1$, and $\xi_{\mathbf{I}_0}(A\to B,C\to D)=0$, the conclusion can be proved similarly to ②.

④ If $\xi_{\mathbf{I}_0}(A,C)=1$, $\xi_{\mathbf{I}_0}(B,D)=0$, and $\xi_{\mathbf{I}_0}(A\to B,C\to D)=1$, then it's clear from (5) that $\xi(A,C)\leq 1,\xi(B,D)\leq\frac{1}{2}$, and $\xi(A\to B,C\to D)\geq\frac{1}{2}$. So we have $\xi(A,C)+\xi(B,D)\leq 1+\frac{1}{2}\leq 1+\xi(A\to B,C\to D)$.

⑤ If $\xi_{\mathbf{I}_0}(A,C)=0$, $\xi_{\mathbf{I}_0}(B,D)=1$, and $\xi_{\mathbf{I}_0}(A\to B,C\to D)=1$, we can prove the conclusion similarly to④.

⑥ If $\xi_{\mathbf{I}_0}(A,C)=0$, $\xi_{\mathbf{I}_0}(B,D)=0$, then we have easily that $\xi(A,C)\leq\frac{1}{2},\xi(B,D)\leq\frac{1}{2}$. So it's clear that $\xi(A,C)+\xi(B,D)\leq\frac{1}{2}+\frac{1}{2}=1\leq 1+\xi(A\to B,C\to D)$.

(iii) It can be easily proved from Proposition 3(iii) and (5).

Corollary 2. *For any $A,B,C\in\mathcal{F}$, if $\xi(A,B)=\xi(B,C)=1$, then $\xi(A,C)=1$.*

3.3 Logically Quasi-equivalent Relation between Formulae

Definition 6. *Suppose that $A,B\in\mathcal{F}$. If $A\to B$ and $B\to A$ are both true formulae in every finite interpretation, we say that A and B are logically quasi-equivalent, denoted by $A\approx_q B$.*

Theorem 7. *Suppose that $A,B\in\mathcal{F}$. If $A\approx_q B$, then $\tau(A)=\tau(B)$; but not vice versa.*

Proof. According to Definition 6, we have that $\|A\|_{\mathbf{I},v}=\|B\|_{\mathbf{I},v}$ holds for every finite interpretation $\mathbf{I}$ and every valuation $v\in\Omega_{\mathbf{I}}$. Thus it follows from (2) and (3) that $\tau(A)=\tau(B)$.

But the inverse proposition doesn't hold. For example: Let $A=P(x)$ and $B=Q(x)$. It can be easily proved that $\tau(A)=\tau(B)=\frac{1}{2}$ from (3), while $A\approx_q B$ doesn't hold.

Remark 4. Even so, if we denote the contradiction by 0 and the logically efficient formula by 1, then for any $A\in\mathcal{F}$, $A\approx_q 0$ if and only if $\tau(A)=0$, and $A\approx_q 1$ if and only if $\tau(A)=1$. In fact, we can prove the following theorem:

Theorem 8. *Suppose that $A,B\in\mathcal{F}$.*
(i) $A\approx_q B$ if and only if $\xi(A,B)=1$;
(ii) $A\approx_q\neg B$ if and only if $\xi(A,B)=0$;
(iii) "$\approx_q$" is an equivalence relation on $\mathcal{F}$.

Remark 5. By the equivalence relation $\approx_q$, we can partition the set $\mathcal{F}$ of formulae into some equivalent classes.In fact,the classification according to the quasi-equivalence relation is not perfect because the logically efficient formulae fall into the same category with the true formulae in any finite interpretations, but we know that the two formulae are different.Even so, according to Theorem 7, this kind of classification is more precise than that by the degrees of the quasi-truth of formulae. Moreover, owing to the density of the distribution of quasi-truth degrees of formulae, the classification is quite accurate.

Remark 6. Furthermore, we can define the operators such as $\vee,\wedge,\to$ and $'$ on these equivalent classes, and discuss the algebraic properties, which are our consequent work.

4 The Pseudo-metric on the Set $\mathcal{F}$ of Formulae

Let $\xi : \mathcal{F} \longrightarrow \mathcal{F}$ be the quasi-similarity degree function defined as (5). Define

$$\rho(A, B) = 1 - \xi(A, B), A, B \in \mathcal{F}. \tag{10}$$

It then follows from Theorem 6(i) that

$$\rho(A, B) + \rho(B, C) \geq \rho(A, C), A, B, C \in \mathcal{F}.$$

Moreover, $\rho(A, A) = 0$ and $\rho(A, B) = \rho(B, A)$ clearly hold, hence ρ is a pseudo-metric on $\mathcal{F}$.

Definition 7. *According to (10), define $\rho : \mathcal{F} \times \mathcal{F} \longrightarrow$[0,1], then ρ is a pseudo-metric on $\mathcal{F}$. $(\mathcal{F}, \rho)$ is called a logic metric space.*

As an immediate consequence of Theorem 1 and Theorem 5, we have:

Theorem 9. *Suppose that $A, B \in \mathcal{F}$.*
(i) $\rho(A, B) = 0$ if and only if $A \approx_q B$; $\rho(A, B) = 1$ if and only if $A \approx_q \neg B$;
(ii) $\rho(A, B) + \rho(A, \neg B) = 1$;
(iii) $\rho(A, \bar{0}) = \tau(A)$, where $\bar{0}$ refers to the set of formulae whose quasi-truth degrees are all equal to 0.

Theorem 10. *Suppose that $A_n, A \in \mathcal{F}$, n=1,2,$\cdots$. If $\lim\limits_{n\to\infty} \rho(A_n, A) = 0$, then $\lim\limits_{n\to\infty} \rho(\neg A_n, \neg A) = 0$.*

Proof. It follows from Theorem 9 that $\rho(\neg A_n, \neg A) = 1 - \rho(\neg A_n, A) = 1 - (1 - \rho(A_n, A)) = \rho(A_n, A)$. So we have that $\lim\limits_{n\to\infty} \rho(\neg A_n, \neg A) = \lim\limits_{n\to\infty} \rho(A_n, A) = 0$.

Theorem 11. *Suppose that $A_n, A, B_n, B \in \mathcal{F}$, n=1,2,$\cdots$. If $\lim\limits_{n\to\infty} \rho(A_n, A) = 0$, $\lim\limits_{n\to\infty} \rho(B_n, B) = 0$, then $\lim\limits_{n\to\infty} \rho(A_n \to B_n, A \to B) = 0$.*

Proof. It follows from (10) and Theorem 6(ii) that

$$\begin{aligned} 0 \leq \rho(A_n \to B_n, A \to B) &= 1 - \xi(A_n \to B_n, A \to B) \\ &\leq 1 - (\xi(A_n, A) + \xi(B_n, B) - 1) \\ &= 1 - (1 - \rho(A_n, A) + 1 - \rho(B_n, B) - 1) \\ &= \rho(A_n, A) + \rho(B_n, B) \longrightarrow 0. \end{aligned}$$

So we have $\lim\limits_{n\to\infty} \rho(A_n \to B_n, A \to B) = 0$.

Corollary 3. *The binary operators $\vee$ and $\wedge$ are both continuous on $(\mathcal{F}, \rho)$.*

Remark 7. In fact, the above theorems tell us that the unary operator $\neg$ and the binary operators $\to$,$\vee$,$\wedge$ are all continuous on $(\mathcal{F}, \rho)$.

5 Conclusion

The elementary theory of quantitative logic include five parts:the truth degrees of formulae, the similarity degree between formulae, logic metric space, theory of approximate reasoning, the degree of consistency of theories. Due to the complexity of the interpretations of the first-order language, we consider the case of truth degree in predicate logic from the viewpoint of all the finite interpretations.Based on the given definition of the quasi-truth degrees of formulae, the degree of the quasi-similarity between formulae is defined, which sequentially induces a pseudo-metric on $\mathcal{F}$. Thus a basic framework for approximate reasoning on $\mathcal{F}$ is provided. In the case of the approximate reasoning on $\mathcal{F}$ and the degree of consistency of theories, we will present our results in another paper.

Acknowledgements. This work described here is partially supported by the grants from the National Natural Science Foundation of China (NSFC No. 60875034).

References

1. Wang, G., Zhou, H.: Quantitative logic. Information Sciences 179, 226–247 (2009)
2. Wang, G.J., Fu, L.: Theory of truth degrees of propositions in two-valued logic. Sci. China A 11, 998–1008 (2001)
3. Wang, G.J., Leung, Y.: Integrated semantics and logic metric spaces. Fuzzy Sets Syst. 136, 71–91 (2003)
4. Wang, G.J., Li, B.J.: Theory of truth degrees of formulas in Łukasiewicz n-valued proportional logic and a limit theorem. Sci. China E 6, 561–569 (2005)
5. Zhou, H.J., Wang, G.J.: A new theory consistency index based on deduction theorems in several logic systems. Fuzzy Sets Syst. 157, 427–443 (2006)
6. Zhou, H.J., Wang, G.J.: Generalized consistency degrees of theories w.r.t. formulas in several standard complete logic systems. Fuzzy Sets Syst. 157, 2058–2073 (2006)
7. Ang, Y.: Ideas and methods of mathematical logic. Fudan University Press, Shanghai (1991) (in Chinese)
8. Wang, G.J., Qin, X.Y., Zhou, X.N.: Theory of quasi-truth degrees of formulas in two-valued predicate logic. Journal of Shaanxi Normal University (Natural Science Edition) 1, 1–6 (2005) (in Chinese)
9. Hamilton, A.G.: Logic for mathematicians. Cambridge University Press, London (1978)
10. Barnes, D.W., Mack, J.M.: An alegebraic introduction to mathematical logic. Springer, New York (1975)
11. Hájek, P.: Metamathematics of fuzzy Logic. Kluwer Academic Publishers, London (1998)
12. Halmos, P.R.: Measure theory. Van Nostrand, New York (1950)
13. Wang, G.J., Qin, X.Y., Zhou, X.N.: An intrinsic fuzzy set on the universe of discourse of predicate formulas. Fuzzy Sets Syst. 157, 3145–3158 (2006)

Conditional Truth Degree of a Logic Theory in Two-Valued Propositional Logic System and Its Application

Xiang-ni Gao and Guo-jun Wang

Institute of Mathematics, Shaanxi Normal University, Xian 710062, China
e-mail: 0310gaoxiangni@163.com, gjwang@snnu.edu.cn

Abstract. The concept of conditional truth degrees of logic theories is proposed in two-valued propositional logic system **L** in the present paper, and this concept is generalized from individual to collective. The concept of divergence degree can be simplified thereby. Moreover, the relation $\eta(\Gamma \mid \Sigma) = \frac{1}{2}(1 + \tau(\Gamma \mid \Sigma))$ between conditional truth degree and conditional consistency degree of a given logic theory Γ w.r.t. Σ is obtained when Γ is consistent. Finally, theories Γ_1 and Γ_2 are divided into six categories, in which the relation of Σ-truth degrees,as well as Σ-consistency degrees and Σ-divergency degrees of logic theories Γ_1, Γ_2 and $\Gamma_1 \cup \Gamma_2$ are compared, respectively.

Keywords: Σ-truth degree of a logic theory, Σ-consistency degree, Σ-divergency degree, finite, countable.

1 Introduction

As is well known,symbolization and formalization are the essential characteristics of mathematical logic,which is quite distinct from computational mathematics. The former lays stress on formal deduction and rigorous argument,while the later concerns with numerical computation and permits approximate solving.In two-valued propositional logic **L**,the concept of truth degree of propositional formulas of mainly characteristics with numerical computation has been proposed in references [1] and [2]. Professor Wang Guojun established the theory of quantitative logic [3] by grading the basic concepts in propositional logics,which was a bridge between artificial intelligence and computational mathematics.In quantitative logic,the concepts of truth degree of formulas was given,moreover,the similarity degree between two formulas and pseudo-metric among formulas were proposed.From then on,there are a series of research results in quantitative logic. However,all

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 285–296.
springerlink.com

these results [4-13] are obtained based only on truth degree of formulas,not considering truth degree among formulas.Based on this, the concept of truth degrees of logic theories is proposed firstly in reference [14],and this concept is generalized from individual to collective.On the basis of truth degrees of logic theories and the theory of conditional probability,the concept of conditional truth degrees of logic theories is introduced in two-valued propositional logic system **L**, which is generalized from individual to collective. The concept of conditional divergency degree can be simplified thereby. Relations between Σ-truth degree, Σ-consistency degree and Σ-divergency degree of a given logic theory is discussed. Moreover, relations of Σ-truth degrees, as well as Σ-consistency degrees and Σ-divergency degrees of logic theories Γ_1, Γ_2 and $\Gamma_1 \cup \Gamma_2$ are compared in two-valued logic system **L**.

2 Preliminaries

This paper will mainly discuss the two-valued propositional logic systems **L**,and the corresponding algebra of which is Boolean-algebra.The lattice is $L = \{0, 1\}$. The implication operator and the corresponding t-norm in which defined as follows:

$$R_B(x, y) = 0 \text{ iff } x = 1, y = 0; \quad x * y = 1 \text{ iff } x = y = 1, \quad x, y \in \{0, 1\}.$$

Let $S = \{p_1, p_2, \cdots\}$ be a countable set, $F(S)$ is the free algebra of type $(\neg, \rightarrow)$ generated by S,where $\neg$ is a unary operator and $\rightarrow$ is binary operator. Elements of $F(S)$ are called propositions or formulas and that of S are called atomic propositions or atomic formulas.

Definition 2.1 [15]. *In logic system* **L**, *let $A = A(p_1, \cdots, p_n)$ be logic formula,then the truth degree of A is defined by*

$$\tau(A) = \frac{|\overline{A}^{-1}|}{2^n};$$

Definition 2.2 [14]. *Let $\Gamma \subseteq F(S)$,then*

$$\tau(\Gamma) = \inf\{\tau(A) \mid A \in D(\Gamma)\}$$

is called the truth degree of the logic theory Γ.

Definition 2.3 [4]. *In logic system* **L**, *let $A \in F(S), \Sigma \subseteq F(S), \Sigma = \{B_1, B_2, \cdots, B_n\}$. Denote $\wedge\Sigma = B_1 \wedge B_2 \wedge \cdots \wedge B_n$. If $\tau(\wedge\Sigma) > 0$, then conditional truth degree of A w.r.t. Σ is defined by*

$$\tau(A \mid \Sigma) = \frac{\tau(A \wedge (\wedge\Sigma))}{\tau(\wedge\Sigma)},$$

Σ-truth degree of A for short.

In the following, suppose that $\Sigma \subseteq F(S)$, $\Sigma = \{B_1, B_2, \cdots, B_n\}$ and $\tau(\wedge\Sigma) > 0$.

Definition 2.4 [4, 14]**.** *In logic system* $\mathbf{L}$*, let* $A, B \in F(S), \Gamma, \Sigma \subseteq F(S)$*,then*

(i) the conditional similarity degree between formulas A *and* B *w.r.t.* Σ *is defined by*

$$\xi_\Sigma(A, B) = \tau((A \to B) \wedge (B \to A) \mid \Sigma),$$

Σ*-similarity degree between* A *and* B *for short;*

(ii) the conditional logical pseudo-distance on $F(S)$ *w.r.t.* Σ *is defined by*

$$\rho_\Sigma(A, B) = 1 - \xi_\Sigma(A, B),$$

Σ*-logical pseudo-distance on* $F(S)$ *for short;*

(iii) the conditional divergency degree of a logic theory Γ *w.r.t.* Σ *is defined by*

$$\mathtt{div}(\Gamma \mid \Sigma) = \sup\{\rho_\Sigma(A, B) \mid A, B \in D(\Gamma)\},$$

Σ*-divergency degree of* Γ *for short;*

(iv) the conditional consistency degree of a logic theory Γ *w.r.t.* Σ *is defined by*

$$\eta(\Gamma \mid \Sigma) = 1 - \frac{1}{2}\mathtt{div}(\Gamma \mid \Sigma)(1 + i(\Gamma)),$$

Σ*-consistency degree of* Γ *for short.*

3 Conditional Truth Degrees of Theories

Definition 3.1. *In logic system* $\mathbf{L}$*, let* $\Gamma, \Sigma \subseteq F(S)$*,then the conditional truth degree of the logic theory* Γ *w.r.t.* Σ *is defined by*

$$\tau(\Gamma \mid \Sigma) = \inf\{\tau(A \mid \Sigma) \mid A \in D(\Gamma)\},$$

Σ*-truth degree of a theory for short.*

Proposition 3.1. *In logic system* $\mathbf{L}$*, let* $\Gamma, \Sigma \subseteq F(S)$*,then the relation between*Σ*-truth degree and* Σ*-divergency degree of the theory* Γ *is given as following*

$$\mathtt{div}(\Gamma \mid \Sigma) = 1 - \tau(\Gamma \mid \Sigma).$$

Proof. On the one hand, suppose that $A, B \in D(\Gamma)$,then $A \wedge B \in D(\Gamma)$.From $\vdash A \wedge B \wedge (\wedge\Sigma) \to (A \to B) \wedge (B \to A) \wedge (\wedge\Sigma)$ we see that

$$\tau(A \wedge B \wedge (\wedge\Sigma)) \leq \tau((A \to B) \wedge (B \to A) \wedge (\wedge\Sigma)). \tag{3.1}$$

Then divide both sides of (3.1) by $\tau(\wedge\Sigma)$ and we have from Definition 3.1 that $\tau(A \wedge B \mid \Sigma) \leq \tau((A \to B) \wedge (B \to A) \mid \Sigma)$, therefore $\rho_\Sigma(A, B) = 1 - \tau((A \to B) \wedge (B \to A) \mid \Sigma) \leq 1 - \tau(A \wedge B \mid \Sigma)$. Thus

$$\begin{aligned}\mathtt{div}(\Gamma \mid \Sigma) &= \sup\{\rho_\Sigma(A,B) \mid A,B \in D(\Gamma)\}\\ &\leq \sup\{1-\tau(A\wedge B \mid \Sigma) \mid A\wedge B \in D(\Gamma)\}\\ &= 1-\inf\{\tau(A\wedge B \mid \Sigma) \mid A\wedge B \in D(\Gamma)\}\\ &\leq 1-\inf\{\tau(C \mid \Sigma) \mid C \in D(\Gamma)\}\\ &= 1-\tau(\Gamma \mid \Sigma).\end{aligned}$$

On the other hand, let $C \in D(\Gamma)$,for any theorem T of $D(\Gamma)$ we have$\rho_\Sigma(T,C) = 1-\tau(C \mid \Sigma)$,therefore

$$\begin{aligned}1-\tau(\Gamma \mid \Sigma) &= 1-\inf\{\tau(C \mid \Sigma) \mid C \in D(\Gamma)\}\\ &= \sup\{1-\tau(C \mid \Sigma) \mid T, C \in D(\Gamma)\}\\ &= \sup\{\rho_\Sigma(T,C) \mid T,C \in D(\Gamma)\}\\ &\leq \sup\{\rho_\Sigma(A,B) \mid A,B \in D(\Gamma)\}\\ &= \mathtt{div}(\Gamma \mid \Sigma).\end{aligned}$$

Hence $\mathrm{div}(\Gamma \mid \Sigma) = 1-\tau(\Gamma \mid \Sigma)$.

Proposition 3.1 points out that Σ-divergent degrees of a theory can skip the intermediate link, which is the Σ-similarity degree of a theory, and directly defined by the Σ-truth degree of a theory.The concept of Σ-divergency degree of a theory can be simplified thereby.

Proposition 3.2. *In logic system* $\mathbf{L}$, *let* $\Gamma, \Sigma \subseteq F(S)$,*then*

(i) $\tau(\Gamma \mid \Sigma) = \tau(A_1 \wedge A_2 \wedge \cdots \wedge A_l \mid \Sigma)$ *when* $\Gamma = \{A_1, A_2, \cdots, A_l\}$ *finite;*

(ii) $\tau(\Gamma \mid \Sigma) = \lim\limits_{l\longrightarrow\infty} \tau(A_1\wedge A_2\wedge\cdots\wedge A_l \mid \Sigma)$ *when* $\Gamma = \{A_1, A_2, \cdots, A_l, \cdots\}$ *is countable.*

Proof. (i) Take any $C \in D(\Gamma)$, then we have $\vdash A_1 \wedge \cdots \wedge A_l \to C$. Therefore $\vdash A_1 \wedge \cdots \wedge A_l \wedge (\wedge\Sigma) \to C \wedge (\wedge\Sigma)$,thus

$$\tau(A_1 \wedge \cdots \wedge A_l \wedge (\wedge\Sigma)) \leq \tau(C \wedge (\wedge\Sigma)), \tag{3.2}$$

Then divide both sides of (3.2) by $\tau(\wedge\Sigma)$ and we have from Definition 3.1 that

$$\tau(A_1 \wedge \cdots \wedge A_l \mid \Sigma) \leq \tau(C \mid \Sigma).$$

Moreover,$A_1 \wedge \cdots \wedge A_l \in D(\Gamma)$,for arbitrariness of C we see that $A_1 \wedge \cdots \wedge A_l$ is the formula in $D(\Gamma)$ with the least Σ-truth degree. Therefore $\inf\{\tau(C \mid \Sigma) \mid C \in D(\Gamma)\} = \tau(A_1 \wedge \cdots \wedge A_l \mid \Sigma)$,that is,

$$\tau(\Gamma \mid \Sigma) = \tau(A_1 \wedge A_2 \wedge \cdots \wedge A_l \mid \Sigma).$$

(ii) Take any $C \in D(\Gamma)$, then $\exists l \in N$ such that $\vdash A_1 \wedge \cdots \wedge A_l \to C$. Therefore $\vdash A_1 \wedge \cdots \wedge A_l \wedge (\wedge\Sigma) \to C \wedge (\wedge\Sigma)$,thus $\tau(A_1 \wedge \cdots \wedge A_l \wedge (\wedge\Sigma)) \leq \tau(C \wedge (\wedge\Sigma))$.Denote $\wedge\overline{\Sigma} = \overline{B}_1 \wedge \cdots \wedge \overline{B}_n$. Since the function $\overline{A}_1 \wedge \cdots \wedge \overline{A}_l \wedge (\wedge\overline{\Sigma})$ is monotonic decreasing and $A_1 \wedge \cdots \wedge A_l \wedge (\wedge\Sigma) \in D(\Gamma)$, we have that $A_1 \wedge \cdots \wedge A_l \wedge (\wedge\Sigma)$ is the smaller formula sequences in $D(\Gamma)$. Thus

$$\inf\{\tau(C \wedge (\wedge\Sigma)) \mid C \in D(\Gamma)\} = \lim_{l\longrightarrow\infty} \tau(A_1 \wedge \cdots \wedge A_l \wedge (\wedge\Sigma)). \tag{3.3}$$

Then divide both sides of (3.3) by $\tau(\wedge\Sigma)$ and we have from Definition 3.1 that

$$\tau(\Gamma \mid \Sigma) = \lim_{l \longrightarrow \infty} \tau(A_1 \wedge \cdots \wedge A_l \mid \Sigma).$$

From Proposition 3.1 and Definition 2.4(iv) we have the following Proportion:

Proposition 3.3. *In logic system* $\mathbf{L}$*, let* $\Gamma, \Sigma \subseteq F(S)$ *and* $i(\Gamma) = 0$*,then*
(i) if $\Gamma = \{A_1, A_2, \cdots, A_l\}$ *is finite,then*

$$\eta(\Gamma \mid \Sigma) = \frac{1}{2}(1 + \tau(\Gamma \mid \Sigma)) = \frac{1}{2}(1 + \tau(A_1 \wedge A_2 \wedge \cdots \wedge A_l \mid \Sigma));$$

(ii) if $\Gamma = \{A_1, A_2, \cdots, A_l, \cdots\}$ *is countable,then*

$$\eta(\Gamma \mid \Sigma) = \frac{1}{2}(1 + \tau(\Gamma \mid \Sigma)) = \frac{1}{2}(1 + \lim_{l \longrightarrow \infty} \tau(A_1 \wedge A_2 \wedge \cdots \wedge A_l \mid \Sigma)).$$

3.1 Relations of Σ-Truth Degrees, as well as Σ-Consistency Degrees and Σ-Divergency Degrees of Logic Theories

Proposition 4.1. *In logic system* $\mathbf{L}$*, let* $\Gamma, \Sigma \subseteq F(S)$*,then*

$$\tau(\Gamma_1 \cup \Gamma_2 \mid \Sigma) \leq \tau(\Gamma_i \mid \Sigma) \leq \tau(\Gamma_1 \cap \Gamma_2 \mid \Sigma), (i = 1, 2).$$

Proof. From Definition 3.1 we see that the more members of theories the less truth degrees of them, therefore Proposition 4.1 holds.

Proposition 4.2. *In logic system* $\mathbf{L}$*, let* $\Gamma, \Sigma \subseteq F(S)$ *and* $i(\Gamma_1) = i(\Gamma_2) = 0$*, then*
(i) $\tau(\Gamma_1 \cup \Gamma_2 \mid \Sigma) \leq \tau(\Gamma_1 \mid \Sigma) + \tau(\Gamma_2 \mid \Sigma)$;
(ii) $\mathrm{div}(\Gamma_1 \cup \Gamma_2 \mid \Sigma) \geq \mathrm{div}(\Gamma_1 \mid \Sigma) + \mathrm{div}(\Gamma_2 \mid \Sigma) - 1$;
(iii) $\eta(\Gamma_1 \cup \Gamma_2 \mid \Sigma) \leq \eta(\Gamma_1 \mid \Sigma) + \eta(\Gamma_2 \mid \Sigma) - \frac{1}{2}$.

Proof. (i) can be directly verified from Proposition 4.1.
(ii) From Proposition 3.1 and Proposition 4.2(i) we can obtain that

$$\begin{aligned} \mathtt{div}(\Gamma_1 \cup \Gamma_2 \mid \Sigma) &= 1 - \tau(\Gamma_1 \cup \Gamma_2 \mid \Sigma) \\ &\geq 1 - \tau(\Gamma_1 \mid \Sigma) - \tau(\Gamma_2 \mid \Sigma) \\ &= 1 - \tau(\Gamma_1 \mid \Sigma) + 1 - \tau(\Gamma_2 \mid \Sigma) - 1 \\ &= \mathtt{div}(\Gamma_1 \mid \Sigma) + \mathtt{div}(\Gamma_2 \mid \Sigma) - 1. \end{aligned}$$

(iii) 1^0 If $i(\Gamma_1 \cup \Gamma_2) = 0$, then it follows from Proposition 3.3 and Proposition 4.2(i) that

$$\begin{aligned} \eta(\Gamma_1 \cup \Gamma_2 \mid \Sigma) &= \tfrac{1}{2}(1 + \tau(\Gamma_1 \cup \Gamma_2 \mid \Sigma)) \\ &\leq \tfrac{1}{2}(1 + \tau(\Gamma_1 \mid \Sigma) + \tau(\Gamma_2 \mid \Sigma)) \\ &= \tfrac{1}{2}(1 + \tau(\Gamma_1 \mid \Sigma)) + \tfrac{1}{2}(1 + \tau(\Gamma_2 \mid \Sigma)) - \tfrac{1}{2} \\ &= \eta(\Gamma_1 \mid \Sigma) + \eta(\Gamma_2 \mid \Sigma) - \tfrac{1}{2}. \end{aligned}$$

2^0 If $i(\Gamma_1 \cup \Gamma_2) = 1$, then $\eta(\Gamma_1 \cup \Gamma_2 \mid \Sigma) = 0$. Since $i(\Gamma_1) = i(\Gamma_2) = 0$, both Γ_1 and Γ_2 are consistency, thus $\frac{1}{2} \leq \eta(\Gamma_i \mid \Sigma) \leq 1 (i = 1, 2)$. Hence $\eta(\Gamma_1 \cup \Gamma_2 \mid \Sigma) = 0 \leq \eta(\Gamma_1 \mid \Sigma) + \eta(\Gamma_2 \mid \Sigma) - \frac{1}{2}$.

Therefore, the conclusion follows from 1^0 and 2^0.

Proposition 4.2 points out that the fuzzy relations of Σ-truth degrees, as well as Σ-consistency degrees and Σ-divergency degrees of logic theories Γ_1, Γ_2 and $\Gamma_1 \cup \Gamma_2$. There is a question whether the exact relations of that can be obtained? In two-valued propositional logic system **L**, for a kind of theories Γ_1 and Γ_2, we see that the exact relations of Σ-truth degrees, as well as Σ-consistency degrees and Σ-divergency degrees of logic theories Γ_1,Γ_2 and $\Gamma_1 \cup \Gamma_2$. In the following contents, we declare that different alphabets or identical alphabet but different suffixes are distinct formulas.

Proposition 4.3. *In logic system* **L***, let* $\Sigma, \Gamma_1, \Gamma_2 \subseteq F(S), r, \alpha, \beta \in [0, 1]$, *and* $\tau(\wedge\Sigma) = r, \tau(\Gamma_1 \mid \Sigma) = \alpha, \tau(\Gamma_2 \mid \Sigma) = \beta$,*then the following statements hold*

(i) if $\Gamma_1 = \{A_1, \cdots, A_{m_1}, B_1, \cdots, B_{m_2}\}$,
$\Gamma_2 = \{A_1, \cdots, A_{m_1}, C_1, \cdots, C_l\}$,*then*

$$\tau(\Gamma_1 \cup \Gamma_2 \mid \Sigma) = \alpha + \beta - \frac{f_1(l)}{r(f_1(l) + g_1(l))}.$$

Where $f_1(l) = |\overline{A \wedge (B \vee C) \wedge (\wedge\Sigma)}^{-1}(1)|, g_1(l) = |\overline{A \wedge (B \vee C) \wedge (\wedge\Sigma)}^{-1}(0)|$, $A = A_1 \wedge \cdots \wedge A_{m_1}, B = B_1 \wedge \cdots \wedge B_{m_2}, C = C_1 \wedge \cdots \wedge C_l$. *In the following,* A, B *and* C *of (ii)–(iv) are the same as (i).*

(ii) If $\Gamma_1 = \{B_1, \cdots, B_{m_2}\}, \Gamma_2 = \{C_1, \cdots, C_l\}$,*then*

$$\tau(\Gamma_1 \cup \Gamma_2 \mid \Sigma) = \alpha + \beta - \frac{f_2(l)}{r(f_2(l) + g_2(l))}.$$

Where $f_2(l) = |\overline{(B \vee C) \wedge (\wedge\Sigma)}^{-1}(1)|, g_2(l) = |\overline{(B \vee C) \wedge (\wedge\Sigma)}^{-1}(0)|$.

(iii) If $\Gamma_1 = \{A_1, \cdots, A_{m_1}, B_1, \cdots, B_{m_2}\}, \Gamma_2 = \{A_1, \cdots, A_{m_1}, C_1, \cdots, C_l, \cdots\}$,*then*

$$\tau(\Gamma_1 \cup \Gamma_2 \mid \Sigma) = \alpha + \beta - \lim_{l \longrightarrow \infty} \frac{f_1(l)}{r(f_1(l) + g_1(l))}.$$

Where $f_1(l)$ *and* $g_1(l)$ *are the same as (i).*

(iv) If $\Gamma_1 = \{B_1, \cdots, B_{m_2}\}, \Gamma_2 = \{C_1, \cdots, C_l, \cdots\}$,*then*

$$\tau(\Gamma_1 \cup \Gamma_2 \mid \Sigma) = \alpha + \beta - \lim_{l \longrightarrow \infty} \frac{f_2(l)}{r(f_2(l) + g_2(l))}.$$

Where $f_2(l)$ *and* $g_2(l)$ *are the same as (ii).*

(v) If $\Gamma_1 = \{A_1, \cdots, A_{m_1}, B_1, B_3, \cdots, B_{2l-1}, \cdots\}$,
$\Gamma_2 = \{A_1, \cdots, A_{m_1}, B_2, B_4, \cdots,$
$B_{2l}, \cdots\}$,*then*

$$\tau(\Gamma_1 \cup \Gamma_2 \mid \Sigma) = \alpha + \beta - \lim_{l \longrightarrow \infty} \frac{f_3(l)}{r(f_3(l) + g_3(l))}.$$

Where $f_3(l) = |\overline{A \wedge (B \vee C) \wedge (\wedge\Sigma)}^{-1}(1)|, g_3(l) = |\overline{A \wedge (B \vee C) \wedge (\wedge\Sigma)}^{-1}(0)|$, $A = A_1 \wedge \cdots \wedge A_{m_1}, B = B_1 \wedge B_3 \wedge \cdots \wedge B_{2l-1}, C = B_2 \wedge B_4 \wedge \cdots \wedge B_{2l}$.

(vi) If $\Gamma_1 = \{B_1, B_3, \cdots, B_{2l-1}, \cdots\}, \Gamma_2 = \{B_2, B_4, \cdots, B_{2l}, \cdots\}$,*then*

$$\tau(\Gamma_1 \cup \Gamma_2 \mid \Sigma) = \alpha + \beta - \lim_{l \longrightarrow \infty} \frac{f_4(l)}{r(f_4(l) + g_4(l))}.$$

Where $f_4(l) = |\overline{(B \vee C) \wedge (\wedge\Sigma)}^{-1}(1)|, g_4(l) = |\overline{(B \vee C) \wedge (\wedge\Sigma)}^{-1}(0)|$, *$B$ and C are the same as (v).*

Proof. We limit ourselves to the proof of (i),(iii) and (v). Without loss of generality,we can assume the following formulas contain the same atomic formula.

(i) Let $A = A_1 \wedge \cdots \wedge A_{m_1}, B = B_1 \wedge \cdots \wedge B_{m_2}, C = C_1 \wedge \cdots \wedge C_l$.It follows from Proposition 3.2 that

$$\tau(\Gamma_1 \mid \Sigma) = \tau(A \mid \Sigma) = \alpha, \quad \tau(\Gamma_2 \mid \Sigma) = \tau(A \mid \Sigma) = \beta.$$

$$\begin{aligned} \tau(\Gamma_1 \cup \Gamma_2 \mid \Sigma) &= \tau((A \wedge B) \wedge (A \wedge C) \mid \Sigma) \\ &= \tau(A \wedge B \mid \Sigma) + \tau(A \wedge C \mid \Sigma) - \tau((A \wedge B) \vee (A \wedge C) \mid \Sigma) \\ &= \alpha + \beta - \tau(A \wedge (B \vee C) \mid \Sigma) \\ &= \alpha + \beta - \tfrac{\tau(A \wedge (B \vee C) \wedge (\wedge\Sigma))}{\tau(\wedge\Sigma)} \\ &= \alpha + \beta - \tfrac{\tau(A \wedge (B \vee C) \wedge (\wedge\Sigma))}{r}. \end{aligned} \tag{4.1}$$

Suppose that $\overline{A \wedge (B \vee C) \wedge (\wedge\Sigma)}$ is $A \wedge (B \vee C) \wedge (\wedge\Sigma)$-induced Boolean function. Equations $\overline{A \wedge (B \vee C) \wedge (\wedge\Sigma)} = 1$ and $\overline{A \wedge (B \vee C) \wedge (\wedge\Sigma)} = 0$ have $f_1(l) = |\overline{A \wedge (B \vee C) \wedge (\wedge\Sigma)}^{-1}(1)|$ and $g_1(l) = |\overline{A \wedge (B \vee C) \wedge (\wedge\Sigma)}^{-1}(0)|$ solutions, respectively. Therefore,

$$\tau(A \wedge (B \vee C) \wedge (\wedge\Sigma)) = \frac{f_1(l)}{f_1(l) + g_1(l)}. \tag{4.2}$$

Hence it follows from (4.1) and (4.2) that

$$\tau(\Gamma_1 \cup \Gamma_2 \mid \Sigma) = \alpha + \beta - \frac{f_1(l)}{r(f_1(l) + g_1(l))}.$$

(iii) Let $A = A_1 \wedge \cdots \wedge A_{m_1}, B = B_1 \wedge \cdots \wedge B_{m_2}, C = C_1 \wedge \cdots \wedge C_l$. It follows from Proposition 3.2 that

$$\tau(\Gamma_1 \mid \Sigma) = \tau(A \mid \Sigma) = \alpha, \quad \tau(\Gamma_2 \mid \Sigma) = \lim_{l \longrightarrow \infty} \tau(A \mid \Sigma) = \beta.$$

$$\begin{aligned}\tau(\Gamma_1\cup\Gamma_2\mid\Sigma) &= \lim_{l\longrightarrow\infty}\tau((A\wedge B)\wedge(A\wedge C)\mid\Sigma)\\ &= \lim_{l\longrightarrow\infty}[\tau(A\wedge B\mid\Sigma)+\tau(A\wedge C\mid\Sigma)-\tau((A\wedge B)\vee(A\wedge C)\mid\Sigma)]\\ &= \tau(A\wedge B\mid\Sigma)+\lim_{l\longrightarrow\infty}\tau(A\wedge C\mid\Sigma)-\lim_{l\longrightarrow\infty}\tau(A\wedge(B\vee C)\mid\Sigma)\\ &= \alpha+\beta-\lim_{l\longrightarrow\infty}\tfrac{\tau(A\wedge(B\vee C)\wedge(\wedge\Sigma))}{\tau(\wedge\Sigma)}\\ &= \alpha+\beta-\lim_{l\longrightarrow\infty}\tfrac{\tau(A\wedge(B\vee C)\wedge(\wedge\Sigma))}{r}.\end{aligned} \tag{4.3}$$

Hence it follows from (4.2) and (4.3) that

$$\tau(\Gamma_1\cup\Gamma_2\mid\Sigma)=\alpha+\beta-\lim_{l\longrightarrow\infty}\frac{f_1(l)}{r(f_1(l)+g_1(l))}.$$

(v) Let $A=A_1\wedge\cdots\wedge A_{m_1}, B=B_1\wedge B_3\wedge\cdots\wedge B_{2l-1}, C=B_2\wedge B_4\wedge\cdots\wedge B_{2l}$.It follows from Proposition 3.2 that

$$\tau(\Gamma_1\mid\Sigma)=\lim_{l\longrightarrow\infty}\tau(A\mid\Sigma)=\alpha,\quad \tau(\Gamma_2\mid\Sigma)=\lim_{l\longrightarrow\infty}\tau(A\mid\Sigma)=\beta.$$

$$\begin{aligned}\tau(\Gamma_1\cup\Gamma_2\mid\Sigma) &= \lim_{l\longrightarrow\infty}\tau((A\wedge B)\wedge(A\wedge C)\mid\Sigma)\\ &= \lim_{l\longrightarrow\infty}[\tau(A\wedge B\mid\Sigma)+\tau(A\wedge C\mid\Sigma)-\tau((A\wedge B)\vee(A\wedge C)\mid\Sigma)]\\ &= \lim_{l\longrightarrow\infty}\tau(A\wedge B\mid\Sigma)+\lim_{l\longrightarrow\infty}\tau(A\wedge C\mid\Sigma)-\lim_{l\longrightarrow\infty}\tau(A\wedge(B\vee C)\mid\Sigma)\\ &= \alpha+\beta-\lim_{l\longrightarrow\infty}\tfrac{\tau(A\wedge(B\vee C)\wedge(\wedge\Sigma))}{\tau(\wedge\Sigma)}\\ &= \alpha+\beta-\lim_{l\longrightarrow\infty}\tfrac{\tau(A\wedge(B\vee C)\wedge(\wedge\Sigma))}{r}\end{aligned} \tag{4.4}$$

Suppose that $\overline{A\wedge(B\vee C)\wedge(\wedge\Sigma)}$ is $A\wedge(B\vee C)\wedge(\wedge\Sigma)$-induced Boolean function. Equations $\overline{A\wedge(B\vee C)\wedge(\wedge\Sigma)}=1$ and $\overline{A\wedge(B\vee C)\wedge(\wedge\Sigma)}=0$ have $f_3(l)=|\overline{A\wedge(B\vee C)\wedge(\wedge\Sigma)}^{-1}(1)|$ and $g_3(l)=|\overline{A\wedge(B\vee C)\wedge(\wedge\Sigma)}^{-1}(0)|$ solutions, respectively. Therefore,

$$\tau(A\wedge(B\vee C)\wedge(\wedge\Sigma))=\frac{f_3(l)}{f_3(l)+g_3(l)}. \tag{4.5}$$

Hence it follows from (4.4) and (4.5) that

$$\tau(\Gamma_1\cup\Gamma_2\mid\Sigma)=\alpha+\beta-\lim_{l\longrightarrow\infty}\frac{f_3(l)}{r(f_3(l)+g_3(l))}.$$

Proposition 4.4. *In logic system* $\mathbf{L}$,$\Sigma,\Gamma_1,\Gamma_2\subseteq F(S),r,\alpha,\beta\in[0,1]$, *and* $\tau(\wedge\Sigma)=r,\mathtt{div}(\Gamma_1\mid\Sigma)=\alpha,\mathtt{div}(\Gamma_2\mid\Sigma)=\beta$, *then the following statements hold*

(i) If $\Gamma_1=\{A_1,\cdots,A_{m_1},B_1,\cdots,B_{m_2}\}$,
$\Gamma_2=\{A_1,\cdots,A_{m_1},C_1,\cdots,C_l\}$,*then*

$$\mathtt{div}(\Gamma_1\cup\Gamma_2\mid\Sigma)=\alpha+\beta-1+\frac{f_1(l)}{r(f_1(l)+g_1(l))}.$$

Where $f_1(l)$ *and* $g_1(l)$ *are the same as Proposition 4.3(i).*
(ii) If $\Gamma_1 = \{B_1, \cdots, B_{m_2}\}$,$\Gamma_2 = \{C_1, \cdots, C_l\}$,*then*

$$\mathtt{div}(\Gamma_1 \cup \Gamma_2 \mid \Sigma) = \alpha + \beta - 1 + \frac{f_2(l)}{r(f_2(l) + g_2(l))}.$$

Where $f_2(l)$ *and* $g_2(l)$ *are the same as Proposition 4.3(ii).*
(iii) If $\Gamma_1 = \{A_1, \cdots, A_{m_1}, B_1, \cdots, B_{m_2}\}$,
$\Gamma_2 = \{A_1, \cdots, A_{m_1}, C_1, \cdots, C_l, \cdots\}$,*then*

$$\mathtt{div}(\Gamma_1 \cup \Gamma_2 \mid \Sigma) = \alpha + \beta - 1 + \lim_{l \longrightarrow \infty} \frac{f_1(l)}{r(f_1(l) + g_1(l))}.$$

Where $g_1(l)$ *are the same as Proposition 4.3(iii).*
(iv) If $\Gamma_1 = \{B_1, \cdots, B_{m_2}\}, \Gamma_2 = \{C_1, \cdots, C_l, \cdots\}$,*then*

$$\mathtt{div}(\Gamma_1 \cup \Gamma_2 \mid \Sigma) = \alpha + \beta - 1 + \lim_{l \longrightarrow \infty} \frac{f_2(l)}{r(f_2(l) + g_2(l))}.$$

Where $f_2(l)$ *and* $g_2(l)$ *are the same as Proposition 4.3(iv).*
(v) If $\Gamma_1 = \{A_1, \cdots, A_{m_1}, B_1, B_3, \cdots, B_{2l-1}, \cdots\}$,
$\Gamma_2 = \{A_1, \cdots, A_{m_1}, B_2, B_4, \cdots,$
$B_{2l}, \cdots\}$,*then*

$$\mathtt{div}(\Gamma_1 \cup \Gamma_2 \mid \Sigma) = \alpha + \beta - 1 + \lim_{l \longrightarrow \infty} \frac{f_3(l)}{r(f_3(l) + g_3(l))}.$$

Where $f_3(l)$ *and* $g_3(l)$ *are the same as Proposition 4.3(v).*
(vi) If $\Gamma_1 = \{B_1, B_3, \cdots, B_{2l-1}, \cdots\}, \Gamma_2 = \{B_2, B_4, \cdots, B_{2l}, \cdots\}$,*then*

$$\mathtt{div}(\Gamma_1 \cup \Gamma_2 \mid \Sigma) = \alpha + \beta - 1 + \lim_{l \longrightarrow \infty} \frac{f_4(l)}{r(f_4(l) + g_4(l))}.$$

Where $f_4(l)$ *and* $g_4(l)$ *are the same as Proposition 4.3(vi).*

Proof. We limit ourselves to the proof of (iii) From Proposition 3.1 and Proposition 4.3(iii) we see that

$$\mathtt{div}(\Gamma_1 \mid \Sigma) = 1 - \tau(\Gamma_1 \mid \Sigma) = \alpha, \quad \mathtt{div}(\Gamma_2 \mid \Sigma) = 1 - \tau(\Gamma_2 \mid \Sigma) = \beta,$$

$$\begin{aligned}\mathtt{div}(\Gamma_1 \cup \Gamma_2 \mid \Sigma) &= 1 - \tau(\Gamma_1 \cup \Gamma_2 \mid \Sigma) \\ &= 1 - \tau(\Gamma_1 \mid \Sigma) - \tau(\Gamma_2 \mid \Sigma) + \lim_{l \longrightarrow \infty} \tfrac{f_1(l)}{r(f_1(l)+g_1(l))} \\ &= \alpha + \beta - 1 + \lim_{l \longrightarrow \infty} \tfrac{f_1(l)}{r(f_1(l)+g_1(l))}.\end{aligned}$$

Proposition 4.5. *In logic system* $\mathbf{L}$*, let* $\Sigma, \Gamma_1, \Gamma_2 \subseteq F(S), r, \alpha, \beta \in [0, 1]$*, and* $\tau(\wedge\Sigma) = r, \eta(\Gamma_1 \mid \Sigma) = \alpha, \eta(\Gamma_2 \mid \Sigma) = \beta$*,* $i(\Gamma_1) = i(\Gamma_2) = i(\Gamma_1 \cup \Gamma_2) = 0$*, then the following statements hold:*

(i) If $\Gamma_1 = \{A_1, \cdots, A_{m_1}, B_1, \cdots, B_{m_2}\}$,
$\Gamma_2 = \{A_1, \cdots, A_{m_1}, C_1, \cdots, C_l\}$,*then*

$$\eta(\Gamma_1 \cup \Gamma_2 \mid \Sigma) = \alpha + \beta - \frac{1}{2} - \frac{f_1(l)}{2r(f_1(l) + g_1(l))}.$$

Where $f_1(l)$ *and* $g_1(l)$ *are the same as Proposition 4.3(i).*
(ii) If $\Gamma_1 = \{B_1, \cdots, B_{m_2}\}$,$\Gamma_2 = \{C_1, \cdots, C_l\}$,*then*

$$\eta(\Gamma_1 \cup \Gamma_2 \mid \Sigma) = \alpha + \beta - \frac{1}{2} - \frac{f_2(l)}{2r(f_2(l) + g_2(l))}.$$

Where $f_2(l)$ *and* $g_2(l)$ *are the same as Proposition 4.3(ii).*
(iii) If $\Gamma_1 = \{A_1, \cdots, A_{m_1}, B_1, \cdots, B_{m_2}\}$,
$\Gamma_2 = \{A_1, \cdots, A_{m_1}, C_1, \cdots, C_l, \cdots\}$,*then*

$$\eta(\Gamma_1 \cup \Gamma_2 \mid \Sigma) = \alpha + \beta - \frac{1}{2} - \lim_{l \longrightarrow \infty} \frac{f_1(l)}{2r(f_1(l) + g_1(l))}.$$

Where $g_1(l)$ *are the same as Proposition 4.3(iii).*
(iv) If $\Gamma_1 = \{B_1, \cdots, B_{m_2}\}$,$\Gamma_2 = \{C_1, \cdots, C_l, \cdots\}$,*then*

$$\eta(\Gamma_1 \cup \Gamma_2 \mid \Sigma) = \alpha + \beta - \frac{1}{2} - \lim_{l \longrightarrow \infty} \frac{f_2(l)}{2r(f_2(l) + g_2(l))}.$$

Where $f_2(l)$ *and* $g_2(l)$ *are the same as Proposition 4.3(iv).*
(v) If $\Gamma_1 = \{A_1, \cdots, A_{m_1}, B_1, B_3, \cdots, B_{2l-1}, \cdots\}$,
$\Gamma_2 = \{A_1, \cdots, A_{m_1}, B_2, B_4, \cdots,$
$B_{2l}, \cdots\}$,*then*

$$\eta(\Gamma_1 \cup \Gamma_2 \mid \Sigma) = \alpha + \beta - \frac{1}{2} - \lim_{l \longrightarrow \infty} \frac{f_3(l)}{2r(f_3(l) + g_3(l))}.$$

Where $f_3(l)$ *and* $g_3(l)$ *are the same as Proposition 4.3(v).*
(vi) If $\Gamma_1 = \{B_1, B_3, \cdots, B_{2l-1}, \cdots\}, \Gamma_2 = \{B_2, B_4, \cdots, B_{2l}, \cdots\}$,*then*

$$\eta(\Gamma_1 \cup \Gamma_2 \mid \Sigma) = \alpha + \beta - \frac{1}{2} - \lim_{l \longrightarrow \infty} \frac{f_4(l)}{2r(f_4(l) + g_4(l))}.$$

Where $f_4(l)$ *and* $g_4(l)$ *are the same as Proposition 4.3(vi).*

Proof. We limit ourselves to the proof of (v). From Proposition 3.3 we see that

$$\eta(\Gamma_1 \mid \Sigma) = \frac{1}{2}(1 + \tau(\Gamma_1 \mid \Sigma)) = \alpha, \quad \eta(\Gamma_2 \mid \Sigma) = \frac{1}{2}(1 + \tau(\Gamma_2 \mid \Sigma)) = \beta,$$

therefore

$$\tau(\Gamma_1 \mid \Sigma) = 2\alpha - 1, \quad \tau(\Gamma_2 \mid \Sigma) = 2\beta - 1.$$

Hence it follows from Proposition 3.3 and Proposition 4.3(v) that

$$\begin{aligned}\eta(\Gamma_1 \cup \Gamma_2 \mid \Sigma) &= \tfrac{1}{2}(1+\tau(\Gamma_1 \cup \Gamma_2 \mid \Sigma)) \\ &= \tfrac{1}{2}(1+\tau(\Gamma_1 \mid \Sigma)+\tau(\Gamma_2 \mid \Sigma) - \lim_{l\longrightarrow\infty} \tfrac{f_3(l)}{r(f_3(l)+g_3(l))}) \\ &= \tfrac{1}{2}(1+2\alpha-1+2\beta-1-\lim_{l\longrightarrow\infty} \tfrac{f_3(l)}{r(f_3(l)+g_3(l))}) \\ &= \alpha+\beta-\tfrac{1}{2}-\lim_{l\longrightarrow\infty} \tfrac{f_3(l)}{2r(f_3(l)+g_3(l))}.\end{aligned}$$

4 Conclusion

In the present paper, the concept of conditional truth degrees of logic theories is introduced in two-valued propositional logic system **L**, which is generalized from individual to collective. The concept of conditional divergency degree can be simplified thereby. Relations between Σ-truth degree, Σ-consistency degree and Σ-divergency degree of a given logic theory is discussed. Finally, relations of Σ-truth degrees, as well as Σ-consistency degrees and Σ-divergency degrees of logic theories Γ_1, Γ_2 and $\Gamma_1 \cup \Gamma_2$ are compared in the logic system **L**. Whether the concept of conditional truth degrees of logic theories is introduced in logic systems $\mathcal{L}^*$,Łukasiewicz and Gödel? Moreover,whether relations of Σ-truth degrees, as well as Σ-consistency degrees and Σ-divergency degrees of logic theories Γ_1, Γ_2 and $\Gamma_1 \cup \Gamma_2$ are compared in these logic systems ? We will study in the future.

Acknowledgements. This work described here is partially supported by the grants from the National Natural Science Foundation of China (NSFC No.10771129).

References

1. Wang, G.J., Fu, L., Song, J.S.: Theory of truth degrees of propositional in two-valued logic. Science in China (Ser. A) 45, 1106–1116 (2002)
2. Wang, G.J., Wang, W.: Logic metric spaces. Acta Mathematics Sinica 44(1), 159–168 (2001)
3. Wang, G.J.: Quantified logic. Chinese J. Engine Math. 23(2), 191–215 (2006) (in Chinese)
4. Hang, B.H., Wang, G.J.: Theory of conditional truth degrees in two-valued propositional logic. Fuzzy Systems and Mathematics 21(4), 9–15 (2007) (in Chinese)
5. Zhang, J.C., Wang, G.J.: Consistency degrees of theories in Łukasiewicz propositional fuzzy logic. Advances in Mathematics 36(6), 761–768 (2007) (in Chinese)
6. Wang, G.J., She, Y.H.: Topological description of consistency of two-valued propositional theories. Acta Mathematica Sinica 50(4), 841–849 (2007) (in Chinese)
7. Wang, G.J., Song, J.S.: Grade method in propositional logic. Acta Electronic Sinica 34(2), 252–257 (2006) (in Chinese)

8. Li, J., Wang, G.: Theory of propositional truth degrees in the logical $\mathcal{L}^*$ system. Science in China (ser. E) 36(6), 631–643 (2006) (in Chinese)
9. Li, B.J., Wang, G.J.: Theory of truth degrees of formulas in Łukasiewicz n-valued propositional logic and a limit theorem. Science in China (ser. E) 35(6), 561–569 (2005) (in Chinese)
10. Li, L.F., Wang, G.J.: Distribution of integrated truth degrees,divergent degrees and consistency degrees in Łukasiewicz logic. Joural of Shaanxi Normal University 34(2), 5–8 (2006) (in Chinese)
11. Zhang, X.F.: Lower truth degree and consistency degrees of theories in five standard complete logic systems. Fuzzy Sets and Systems 21(6), 2058–2073 (2007) (in Chinese)
12. Wang, G.J., Hui, X.J.: Randomization of classical inference patterns and its application. Science in China (Ser. F) 50(6), 867–877 (2007)
13. Li, J., Wang, G.J.: Theory of truth degree of propositions in logic system $\mathcal{L}_n^*$. Science in China (Ser. E) 36(6), 631–643 (2006)
14. Wang, J., Gao, X.N.: Truth degree of a logic theory in propositional logic systems and its application. Joural of Shaanxi Normal University 37(1), 1–6 (2009) (in Chinese)
15. Wang, G.J.: Introduction to Mathematical Logic and Resolution Principle. Science in China Press, Beijing (2009)
16. Wang, G.J.: Theory of Non-Classical Mathematical Logic and Approximate Reasoning. Science in China Press, Beijing (2008) (in Chinese)
17. Gottwald, S., NovaK, V.: On the consistency of fuzzy theories. In: Proc. 7th IFSA Word Congr. Academia, Prague, pp. 168–171 (1997)
18. Hajek, P.: Metamathematics of Fuzzy Logic. Kluwer Academic Publishers, Dordrecht (1998)
19. Pei, D.W.: The operation $\otimes$ and deductive theorem in the formal deductive system $\mathcal{L}^*$. Fuzzy Sets and Systems 15(1), 34–39 (2001) (in Chinese)

$PIMP$-Filters of R_0-Algebras

Yong-lin Liu, Mei-ying Ren, and Qiu-sun Ye

Department of Mathematics and Computer, Wuyi University, Wuyishan 354300, Fujian, China
e-mail: ylliu9@126.com

Abstract. The notion of positive implication MP-filters (briefly, $PIMP$-filters) in R_0-algebras is introduced. The characteristic properties and extension property of $PIMP$-filters are obtained. Finally, the positive implicative R_0-algebra is completely described by its $PIMP$-filters.

Keywords: $PIMP$-filter, positive implication R_0-algebra.

1 Introduction

In recent years, motivated by both theory and application, the study of t-norm-based logic systems and the corresponding pseudo-logic systems has been become a greater focus in the field of logic (cf. [2]-[15]). Here, t-norm-based logical investigations were first to the corresponding algebraic investigations, and in the case of pseudo-logic systems, algebraic development was first to the corresponding logical development. The notion of NM-algebras was introduced by Esteva and Godo [3] from the views of the left-continuous t-norms and their residua. In [15], Wang proposed the notion of R_0-algebras. Pei [14] proved that R_0-algebras and NM-algebras are the same algebraic structures. In [8], Liu et al. introduced the notion of positive implication R_0-algebras. In this paper, the notion of positive implication MP-filters (briefly, $PIMP$-filters) in R_0-algebras is introduced. The characteristic properties and extension property of $PIMP$-filters are obtained. Finally, the positive implicative R_0-algebra is completely described by its $PIMP$-filters.

2 Preliminaries

By an R_0-algebra is meant a bounded distributive lattice $(M, \vee, \wedge, 0, 1)$ with order-reversing involution " $'$ " and a binary operation "$\rightarrow$" satisfying the following axioms:

(R_1) $a' \rightarrow b' = b \rightarrow a$,
(R_2) $1 \rightarrow a = a$,
(R_3) $b \rightarrow c \leq (a \rightarrow b) \rightarrow (a \rightarrow c)$,

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 297–302.
springerlink.com

(R_4) $a \to (b \to c) = b \to (a \to c)$,
(R_5) $a \to (b \vee c) = (a \to b) \vee (a \to c)$,
(R_6) $(a \to b) \vee ((a \to b) \to (a' \vee b)) = 1$,
for all $a, b, c \in M$.

In an R_0-algebra, the following hold:

(1) $0 \to a = 1, a \to 0 = a', a \to a = 1$ and $a \to 1 = 1$,
(2) $a \leq b$ implies $b \to c \leq a \to c$ and $c \to a \leq c \to b$,
(3) $a \to b \leq (b \to c) \to (a \to c)$,
(4) $((a \to b) \to b) \to b = a \to b$,
(5) $a \to (b \wedge c) = (a \to b) \wedge (a \to c)$,
(6) $(a \vee b) \to c = (a \to c) \wedge (b \to c)$,
(7) $(a \wedge b) \to c = (a \to c) \vee (b \to c)$,
(8) $a \leq b$ if and only if $a \to b = 1$.

A subset F of an R_0-algebra M is called an MP-filter of M if it satisfies

(F_1) $1 \in F$,
(F_2) $x \in F$ and $x \to y \in F$ imply $y \in F$ for all $x, y \in M$.

An R_0-algebra M is called a positive implication R_0-algebra [8] if it satisfies for all $x, y \in M$,

$$x \to (y \to z) \leq (x \to y) \to (x \to z).$$

3 *PIMP*-Filters

Definition 1. *A subset F of an R_0-algebra M is said to be a positive implication MP-filter (briefly, PIMP-filter) of M if it satisfies (F_1) and*
(F_3) $x \to (y \to z) \in F$ and $x \to y \in F$ imply $x \to z \in F$ for all $x, y, z \in M$.

The relation between $PIMP$-filters and MP-filters in an R_0-algebra is as follows:

Proposition 1. *A PIMP-filter is an MP-filter, but the converse is not true.*

Proof. Assume that F is a $PIMP$-filter. If $x \in F$ and $x \to y \in F$, then $1 \to x \in F$ and $1 \to (x \to y) \in F$. By (F_3), $1 \to y = y \in F$. Hence (F_2) holds. Combining with (F_1), F is an MP-filter. The last part is shown by the following example.

Example 1. Let $\overline{W} = [0, 1]$. For any $a, b \in [0, 1]$, define $a' = 1 - a, a \vee b = max\{a, b\}, a \wedge b = min\{a, b\}$ and

$$a \to b = \begin{cases} 1, & a \leq b, \\ a' \vee b, & \text{otherwise}. \end{cases}$$

Then $\overline{W}$ is an R_0-algebra, which is called R_0 unit interval [15]. $F = \{1\}$ is an MP-filter of $\overline{W}$, but is not a $PIMP$-filter because: $0.3 \to (0.6 \to 0.2) = 1 \in \{1\}$ and $0.3 \to 0.6 = 1 \in \{1\}$, but $0.3 \to 0.2 = 0.7 \notin \{1\}$. The proof is complete.

Next, we investigate the characterizations of $PIMP$-filters in R_0-algebras.

Theorem 1. *If F is an MP-filter of an R_0-algebra M, then the following are equivalent:*

(i) *F is a PIMP-filter of M,*

(ii) *$x \to (x \to y) \in F$ implies $x \to y \in F$ for all $x, y \in M$,*

(iii) *$x \to (y \to z) \in F$ implies $(x \to y) \to (x \to z) \in F$ for all $x, y, z \in M$,*

(iv) *$u \to (x \to (y \to z)) \in F$ and $u \in F$ imply $(x \to y) \to (x \to z) \in F$ for all $x, y, z, u \in M$,*

(v) *$z \to (x \to (x \to y)) \in F$ and $z \in F$ imply $x \to y \in F$ for all $x, y, z \in M$.*

Proof. (i)⇒(ii). Let F be a MP-filter of an R_0-algebra M and $x \to (x \to y) \in F$. Since $x \to x = 1 \in F$, by (F_3) we have $x \to y \in F$.

(ii)⇒(iii). Let $x \to (y \to z) \in F$. Since $x \to (x \to ((x \to y) \to z)) = x \to ((x \to y) \to (x \to z)) \geq x \to (y \to z) \in F$, we have $x \to (x \to ((x \to y) \to z)) \in F$. By (ii) $x \to ((x \to y) \to z) \in F$, i.e., $(x \to y) \to (x \to z) \in F$.

(iii)⇒(iv). Trivial.

(iv)⇒(v). If $z \to (x \to (x \to y)) \in F$ and $z \in F$, by (iv) we have $(x \to x) \to (x \to y) \in F$. That is $x \to y \in F$.

(v)⇒(i). Let $x \to (y \to z) \in F$ and $x \to y \in F$. Since $(x \to y) \to (x \to (x \to z)) \geq y \to (x \to z) = x \to (y \to z) \in F$, $(x \to y) \to (x \to (x \to z)) \in F$. It follows from (v) that $x \to z \in F$. The proof is complete.

Theorem 2. *Let F be an MP-filter of an R_0-algebra M. Then F is a PIMP-filter of M if and only if for any $t \in M$, the subset $F_t = \{x \in M : t \to x \in F\}$ is a MP-filter of M.*

Proof. Assume that for any $t \in M$, F_t is an MP-filter of M. Let $x \to (x \to y) \in F$. Then $x \to y \in F_x$. Since $x \in F_x$ we have $y \in F_x$, and so $x \to y \in F$. By Theorem 1 (ii), F is a $PIMP$-filter of M.

Conversely, let F is a $PIMP$-filter of M and $x \in F_t$, $x \to y \in F_t$. Then $t \to x \in F$ and $t \to (x \to y) \in F$. Since $(t \to (x \to y)) \to (t \to (t \to y)) = (x \to (t \to y)) \to (t \to (t \to y)) \geq t \to x$, we obtain $t \to (t \to y) \in F$. By Theorem 1 (ii), $t \to y \in F$, and so $y \in F_t$. Hence F_t is an MP-filter of M, completing the proof.

Corollary 1. *Let F be a PIMP-filter of an R_0-algebra M. For any $t \in M$, $F_t = \{x \in M : t \to x \in F\}$ is the least MP-filter of M containing F and t.*

Proof. By Theorem 2, F_t is an MP-filter of M. Clearly $F \subseteq F_t$ and $t \in F_t$. If H is an MP-filter containing F and t, then for any $x \in F_t$ we have $t \to x \in F \subseteq H$. It follows that $x \in H$ as $t \in H$. Hence $F_t \subseteq H$, completing the proof.

Corollary 2. *Let M be an R_0-algebra such that $\{1\}$ is a PIMP-filter. For any $a \in M$, $U_a = \{x \in M : a \leq x\}$ is the least MP-filter of M containing a.*

The extension property of $PIMP$-filters in an R_0-algebra is given by the following:

Proposition 2. *Let F and H be two MP-filters of an R_0-algebra M with $F \subseteq H$. If F is a PIMP-filter of M, then so is H.*

Proof. Suppose that F is a $PIMP$-filter of M and $x \to (x \to y) \in H$ for all $x, y \in M$. Putting $t = x \to (x \to y)$, then $x \to (x \to (t \to y)) = t \to (x \to (x \to y)) = 1 \in F$. By Theorem 1 (ii), $x \to (t \to y) \in F \subseteq H$. That is $t \to (x \to y) \in H$. Thus $x \to y \in H$ as H is an MP-filter. Hence H is a $PIMP$-filter of M. This completes the proof.

Proposition 3 ([8]). *If M is an R_0-algebra, then the following are equivalent:*

(i) *M is a positive implication R_0-algebra,*
(ii) $x \to y \geq x \to (x \to y)$,
(iii) $x \to (y \to z) \leq (x \to y) \to (x \to z)$,
(iv) $x \to (y \to z) = (x \to y) \to (x \to z)$.

Proposition 4 ([8]). *Let M be an R_0-algebra. Then the following are equivalent:*

(i) *M is a positive implication R_0-algebra,*
(ii) *$x \leq u$ implies $u \to (x \to y) = x \to y$,*
(iii) *$x \leq u$ implies $u \to (y \to z) = (x \to y) \to (x \to z)$,*
(iv) *$x \leq y \to z$ implies $x \to y \leq x \to z$,*
(v) *$x \leq x \to y$ implies $x \leq y$.*

Finally, we characterize the positive implication R_0-algebras by its $PIMP$-filters.

Theorem 3. *Let M be an R_0-algebra. The following are equivalent:*

(i) *M is a positive implication R_0-algebra,*
(ii) *every MP-filter of M is a PIMP-filter,*
(iii) *The unit MP-filter $\{1\}$ of M is a PIMP-filter,*
(iv) *for all $t \in M$, $U_t = \{x \in M : t \leq x\}$ is a MP-filter,*
(v) *for all MP-filter F of M and all $t \in M$, $F_t = \{x \in M : t \to x \in F\}$ is an MP-filter.*

Proof. (i)⇒(ii). It follows directly from Proposition 3 (iii) and Theorem 1 (ii).

(ii)⇒(iii). Trivial.

(iii)⇒(iv). For any $t \in M$ and $x \to y \in U_t$, $x \in U_t$, we have $t \to x = 1 \in \{1\}$ and $t \to (x \to y) = 1 \in \{1\}$. Analogous to the proof of Theorem 2, we can obtain $t \to (t \to y) = 1 \in \{1\}$. By Theorem 1 (ii) $t \to y \in \{1\}$, which means that $y \in U_t$. Hence U_t is an MP-filter.

(iv)⇒(i). Let $x \leq x \to y$. Then $x \to y \in U_x$. Since U_x is an MP-filter and $x \in U_x$, we have $y \in U_x$. That is $x \leq y$. By Proposition 4 (iv), M is a positive implication R_0-algebra.

(i)⇒(v). For any MP-filter F of M and all $t \in M$, it follows from M is a positive implication R_0-algebra and (ii) that F is a $PIMP$-filter of M. By Theorem 2, F_t is an MP-filter of M.

(v)⇒(iii). Let $x \to (x \to y) \in \{1\}$ and $B = \{u \in M : x \to u \in \{1\}\}$. By the hypothesis B is an MP-filter. Since $x \to y \in B$ and $x \in B$ we have $y \in B$, i.e., $x \to y \in \{1\}$. Hence $\{1\}$ is a $PIMP$-filter as Theorem 1 (ii). The proof is complete.

Let F be a MP-filter of an R_0-algebra M. For any $x, y \in M$, define a binary relation $\sim$ on M by

$$x \sim y \text{ if and only if } x \to y \in F \text{ and } y \to x \in F.$$

Then $\sim$ is a congruence relation on M. Let $C_x = \{y \in M \mid y \sim x\}$ and $M/F = \{C_x \mid x \in M\}$. Then $(M/F; \vee; \wedge; \to; \,'\,; C_0; C_1)$ is a quotient R_0-algebra, where

$$C_x \vee C_y = C_{x\vee y}, \quad C_x \wedge C_y = C_{x\wedge y}, \quad C_x \to C_y = C_{x\to y}, \quad (C_x)' = C_{x'}.$$

Corollary 3. *Let F be an MP-filter of an R_0-algebra M. Then F is a PIMP-filter if and only if M/F is a positive implication R_0-algebra.*

Proof. Suppose that F is a $PIMP$-filter of M. Now we show that unit MP-filter $\{C_1\}$ of M/F is a $PIMP$-filter. If $C_x \to (C_x \to C_y) \in \{C_1\}$, i.e., $C_{x\to(x\to y)} = C_1$. Hence $1 \to (x \to (x \to y)) \in F$, i.e., $x \to (x \to y) \in F$. By Theorem 1 (ii) $x \to y \in F$, i.e., $1 \to (x \to y) \in F$. On the other hand, $(x \to y) \to 1 = 1 \in F$. Hence $C_{x\to y} = C_1$, i.e., $C_x \to C_y \in \{C_1\}$. Thus M/F is a positive implication R_0-algebra by Theorem 3 (iii).

Conversely, if M/F is a positive implication R_0-algebra, by Theorem 3 (iii) $\{C_1\}$ is a $PIMP$-filter. Let $x \to (x \to y) \in F$, i.e., $1 \to (x \to (x \to y)) \in F$. Since $(x \to (x \to y)) \to 1 \in F$, we have $C_{x\to(x\to y)} = C_1 \in \{C_1\}$, i.e., $C_x \to ((C_x \to C_y) \in \{C_1\}$. Hence $C_x \to C_y \in \{C_1\}$. It means that $x \to y \in F$. Therefore F is a $PIMP$-filter of M. The proof is complete.

Summarizing Theorem 3, Propositions 3, 4 and Corollary 3, we have the following corollary.

Corollary 4. *Let M be an R_0-algebra. Then the following are equivalent:*

(1°) *M is a positive implication R_0-algebra,*
(2°) $x \to y \geq x \to (x \to y)$,
(3°) $x \to (y \to z) \leq (x \to y) \to (x \to z)$,
(4°) $x \to (y \to z) = (x \to y) \to (x \to z)$,
(5°) $x \leq u$ *implies* $u \to (x \to y) = x \to y$,
(6°) $x \leq u$ *implies* $u \to (y \to z) = (x \to y) \to (x \to z)$,
(7°) $x \leq y \to z$ *implies* $x \to y \leq x \to z$,
(8°) $x \leq x \to y$ *implies* $x \leq y$,

(9°) *every MP-filter of M is a PIMP-filter,*
(10°) *The unit MP-filter* $\{1\}$ *of M is a PIMP-filter,*
(11°) *for all* $t \in M$, $U_t = \{x \in M : t \leq x\}$ *is an MP-filter,*
(12°) *for all MP-filter F of M and all* $t \in M$, $F_t = \{x \in M : t \to x \in F\}$ *is an MP-filter,*
(13°) *for all MP-filter F of M, M/F is a positive implication* R_0*-algebra.*

Acknowledgements. This work described here is partially supported by the National Natural Science Foundation of China (Grant no. 60875034, 60775038) and the Science and Technology Foundation of Fujian Education Department (Grant no. JA09242).

References

1. Balbes, R., Dwinger, P.: Distributive lattices. Univ. of Missouri Press, Columbia (1974)
2. Cignoli, R., D'Ottaviano, I.M.L., Mundici, D.: Algebraic foundations of many-valued reasoning. Kluwer Academic Publishers, Dordrecht (2000)
3. Esteva, F., Godo, J.: Monoidal t-norm-based logic: Towards a logic for left-continuous t-norms. Fuzzy Sets and Systems 124, 271–288 (2001)
4. Flondor, P., Georgescu, G., Iorgulescu, A.: Pseudo-t-norms and pseudo-BL algebras. Soft Computing 5(5), 355–371 (2001)
5. Georgescu, G., Iorgulescu, A.: Pseudo-BL algebras: A noncommutative extension of BL algebras. In: Abstracts of The Fifth International Conference FSTA, Slovakia (2000)
6. Hajek, P.: Fleas and fuzzy logic. J. of Mult.-Valued Logic & Soft Computing 11(1-2), 137–152 (2005)
7. Iorgulescu, A.: Some direct ascendents of Wajsberg and MV algebras. Scientiae Mathematicae Japonicae 57, 583–647 (2003)
8. Liu, Y.L., Ren, M.Y., Akram, M.: Positive implication R0-algebras. In: 2009 World Congress on Computer Science and Information Engineering, Los Angeles, California, USA, March 31-April 2 (2009)
9. Liu, Y.L., Meng, J.: Fuzzy ideals in BCI-algebras. Fuzzy Sets and Systems 123, 227–237 (2001)
10. Liu, Y.L., Xu, Y., Meng, J.: BCI-implicative ideals of BCI-algebras. Information Sciences 177, 4987–4996 (2007)
11. Liu, Y.L., Xu, Y., Qin, K.Y., Liu, S.Y.: ILI-ideals and prime LI-ideals in lattice implication algebras. Information Sciences 155, 157–175 (2003)
12. Liu, Y.L., Liu, S.Y., Xu, Y.: Pseudo-BCK algebras and PD-posets. Soft Computing 11(1), 91–101 (2007)
13. Liu, Y.L., Liu, S.Y., Xu, Y.: An answer to the Shim-Lele's open problem on the fuzzy filters. J. Appl. Math. Computing 21, 325–329 (2006)
14. Pei, D.W.: On equivalent forms of fuzzy logic systems NM and $IMTL$. Fuzzy Sets and Systems 138, 187–195 (2003)
15. Wang, G.J.: Non-classical mathematical logic and approximate reasoning. Science Press, Beijing (2000)

Quantitative Logic Theory in Gödel System

Wen-bin Xu

Department of Mathematics, School of Science, Jimei University, Xiamen 361021, Fujian Province, P.R. China
e-mail: wbxu@jmu.edu.cn

Abstract. The purpose of this paper is to establish quantitative logic theory in Gödel system. Some Results about truth degree and resemblance degree are obtained and it is showed that the existing quantitative logic theory also apply to Gödel system through modifying negation operator.

Keywords: Gödel system, truth degree, resemblance degree, pseudo-metric.

1 Introduction

At present, the study of formula truth degree based on the grading idea has been a hot topic in some common logic systems, such as classical two-valued propositional logic[1, 2, 3], many-valued propositional logic[4, 5, 6], predicate logic[7], fuzzy propositional logic[8, 2], model logic[9]. Moreover, based on the truth degree of formula, many new concepts, such as consistent degree of theory Γ[10, 11, 12], resemblance degree between two formulae[13], are proposed. Quantitative logic proposed in [14] based on these results initiate a new theory about fuzzy logic.

The relative study[15, 16] of Gödel logic have been done and some good results have been developed. However, [2] points out that Gödel system are not suitable to establish fuzzy logic based on the strong negation operator. The present paper try to establish quantitative logic theory in Gödel system through modifying negation operator.

2 Preliminaries

Let $S = \{p_1, p_2, \cdots\}$ be a countable set and $\neg$ and $\rightarrow$ be unary and binary logic connectives respectively. Write $F(S)$ be the free algebra of type $(\neg, \rightarrow)$ generated by S.Elements of $F(S)$ are called propositions or formulae and that of S are called atomic propositions or atomic formulas.

Define a binary operator $\rightarrow$ on $[0, 1]$ as follows:

$$a \rightarrow b = \begin{cases} 1, & a \le b \\ b, & a > b \end{cases} \tag{1}$$

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 303–309.
springerlink.com

Then $\rightarrow$ is called Gödel implication operator. Otherwise, define a unary operator $\neg$ and two binary ones $\vee$ and $\wedge$ on $[0,1]$ as follows:

$$\neg a = 1 - a,\ a \vee b = \max\{a, b\},\ a \wedge b = \min\{a, b\}. \tag{2}$$

Obviously,

$$a \vee b = \neg(\neg a \wedge \neg b),\ a \wedge b = \neg(\neg a \vee \neg b).$$

A mapping v from $F(S)$ to unit interval $[0,1]$ is said to be a valuation if v is a homomorphism of type $(\neg, \vee, \rightarrow)$, i.e.,$v(\neg A) = \neg v(A), v(A \vee B) = v(A) \vee v(B), v(A \rightarrow B) = v(A) \rightarrow v(B)$. The set of all valuations of $F(S)$ is denoted by Ω. A formula A is called a tautology, denoted by $\models A$, if $v(A) = 1$ for all valuations v. Conversely, it is called a contradiction if $v(A) = 0$ for all valuations v. A and B are said to be logic equivalent, denoted by $A \approx B$, if $v(A) = v(B)$ holds for all valuation v.

Suppose that $A(p_1, p_2, \cdots, p_m) \in F(S)$ and $p_1, p_2, \cdots, p_m$ are atomic formulas included in A. Then a McNaughton function[17] $\overline{A} : [0,1]^m \rightarrow [0,1]$ induced by A is defined as follows: Substitute x_{i_k} for p_{i_k} in $A(k = 1, \cdots, m)$ and keep the logic connectives in A unchanged but explain them as (1) and (2).

Let $v \in \Omega$ and $A(p_1, p_2, \cdots, p_m) \in F(S)$, then $v(A(p_1, p_2, \cdots, p_m)) = \overline{A}(v(p_1), v(p_2), \cdots, v(p_m))$.

3 Truth Degree of Formulas in Gödel Logic System

[2, 8] gives the definition of integral truth degree in continue valued logic system. We apply it to Gödel system as follows.

Definition 1. *Let $A(p_1, p_2, \cdots, p_m)$ be a formula composed of m atomic propositions. Define*

$$\tau(A) = \int_{[0,1]^m} \overline{A}(x_1, x_2, \cdots, x_m) dx_1 dx_2 \cdots dx_m. \tag{3}$$

Then $\tau(A)$ is called the truth degree *of formula A.*

Theorem 1. *Suppose that $A, B \in F(S)$ in Gödel logic system, then we have the following properties.*

(1) *If A is a tautology, then $\tau(A) = 1$;*
(2) *If A is a contradiction, then $\tau(A) = 0$.*
(3) *If $A \approx B$, then $\tau(A) = \tau(B)$.*
(4) $\tau(\neg A) = 1 - \tau(A)$.
(5) *If A is a tautology, then $\tau(A \rightarrow B) = \tau(B), \tau(B \rightarrow A) = 1$.*
(6) $\tau(A \vee B) + \tau(A \wedge B) = \tau(A) + \tau(B)$.

Proof. (6) is proved as follows, others can be verified easily.

It is easy to verify $a \vee b + a \wedge b = a + b$ for all $a, b \in [0,1]$. Suppose that both A and B consist of same atomic formulas $p_1, p_2, \cdots, p_m$. Then

$$\begin{aligned}
\tau(A \vee B) + \tau(A \wedge B) &= \int_{[0,1]^m} \overline{A \vee B}(x_1, \cdots, x_m) dx_1 \cdots dx_m \\
&\quad + \int_{[0,1]^m} \overline{A \wedge B}(x_1, \cdots, x_m) dx_1 \cdots dx_m \\
&= \int_{[0,1]^m} [\overline{A \vee B}(x_1, \cdots, x_m) \\
&\quad + \overline{A \wedge B}(x_1, \cdots, x_m)] dx_1 \cdots dx_m \\
&= \int_{[0,1]^m} [\overline{A}(x_1, \cdots, x_m) + \overline{B}(x_1, \cdots, x_m)] dx_1 \cdots dx_m \\
&= \int_{[0,1]^m} \overline{A}(x_1, \cdots, x_m) dx_1 \cdots dx_m \\
&\quad + \int_{[0,1]^m} \overline{B}(x_1, \cdots, x_m) dx_1 \cdots dx_m \\
&= \tau(A) + \tau(B).
\end{aligned}$$

Remark 1. Theorem 1(1)through (3) are not sufficient. The counter-example for (1) is given as follows.

Consider the formula $A = ((p_1 \to p_2) \to p_2) \vee ((p_2 \to p_1) \to p_1)$. It is easy to obtain the function $\overline{A}$ induced by A.

$$\overline{A}(x_1, x_2) = \begin{cases} x_1, & x_1 = x_2 \\ 1, & \text{others.} \end{cases}$$

Obviously, $\tau(A) = \int_0^1 \int_0^1 \overline{A}(x_1, x_2) dx_1 dx_2 = 1$, while A is not a tautology.

Theorem 2. *Suppose that $A, B \in F(S)$ in Gödel logic system and $\alpha, \beta \in [0,1]$, then we have the following properties.*

(1)(MP-rule for truth degree) *If $\tau(A) \geq \alpha, \tau(A \to B) \geq \beta$, then $\tau(B) \geq (\alpha + \beta - 1) \vee 0$.*

(2)(HS-rule for truth degree) *If $\tau(A \to B) \geq \alpha, \tau(B \to C) \geq \beta$, then $\tau(A \to C) \geq (\alpha + \beta - 1) \vee 0$.*

Proof. It is easy to verify $b \geq a + (a \to b) - 1$ and $a \to c \geq (a \to b) + (b \to c) - 1$ for all $a, b, c \in [0,1]$. Suppose that A, B and C consist of same atomic formulas $p_1, p_2, \cdots, p_m$. Then

$$\begin{aligned}
\tau(B) &= \int_{[0,1]^m} \overline{B}(x_1, \cdots, x_m) dx_1 \cdots dx_m \\
&\geq \int_{[0,1]^m} [\overline{A}(x_1, \cdots, x_m) \\
&\quad + \overline{A}(x_1, \cdots, x_m) \to \overline{B}(x_1, \cdots, x_m) - 1] dx_1 \cdots dx_m \\
&= \int_{[0,1]^m} \overline{A}(x_1, \cdots, x_m) dx_1 \cdots dx_m \\
&\quad + \int_{[0,1]^m} \overline{A \to B}(x_1, \cdots, x_m) dx_1 \cdots dx_m - \int_{[0,1]^m} 1 dx_1 \cdots dx_m \\
&= \tau(A) + \tau(A \to B) - 1 \\
&\geq \alpha + \beta - 1.
\end{aligned}$$

Thus (1) holds.

$$
\begin{aligned}
\tau(A \to C) &= \int_{[0,1]^m} \overline{A \to C}(x_1, \cdots, x_m) dx_1 \cdots dx_m \\
&= \int_{[0,1]^m} [\overline{A}(x_1, \cdots, x_m) \to \overline{C}(x_1, \cdots, x_m)] dx_1 \cdots dx_m \\
&\geq \int_{[0,1]^m} [\overline{A}(x_1, \cdots, x_m) \to \overline{B}(x_1, \cdots, x_m) \\
&\quad + \overline{B}(x_1, \cdots, x_m) \to \overline{C}(x_1, \cdots, x_m) - 1] dx_1 \cdots dx_m \\
&= \int_{[0,1]^m} \overline{A \to B}(x_1, \cdots, x_m) dx_1 \cdots dx_m \\
&\quad + \int_{[0,1]^m} \overline{B \to C}(x_1, \cdots, x_m) dx_1 \cdots dx_m - \int_{[0,1]^m} 1 dx_1 \cdots dx_m \\
&= \tau(A \to B) + \tau(B \to C) - 1 \\
&\geq \alpha + \beta - 1.
\end{aligned}
$$

Thus (2) holds.

Corollary 1. *Suppose that $A, B, C \in F(S)$, then we have the following properties.*

(1) $\tau(A) + \tau(A \to B) \leq \tau(B) + 1$.
(2) $\tau(A \to B) + \tau(B \to C) \leq \tau(A \to C) + 1$.
(3) *If* $\tau(A) = 1, \tau(A \to B) = 1$, *then* $\tau(B) = 1$.
(4) *If* $\tau(A \to B) = 1, \tau(B \to C) = 1$, *then* $\tau(A \to C) = 1$.
(5) *If* $\models A \to B$, *then* $\tau(A) \leq \tau(B)$.

4 Resemblance Degree between Formulas in Gödel Logic System

[2, 8] gives the definition of resemblance degree in some logic systems. We apply it to Gödel system as follows.

Definition 2. *Suppose that $A, B \in F(S)$. Define*

$$\xi(A, B) = \tau((A \to B) \wedge (B \to A)). \tag{4}$$

Then $\xi(A, B)$ is called the resemblance degree *between Formulas A and B.*

Theorem 3. *Suppose that $A, B, C \in F(S)$ in Gödel logic system, then we have the following properties.*

(1) *If A is a tautology, then* $\xi(A, B) = \tau(B)$.
(2) *If one of A and B is a tautology and other is a contradiction, then* $\xi(A, B) = 0$.
(3) *If* $A \approx B$, *then* $\xi(A, B) = 1$.
(4) $\xi(A, B) + \xi(B, C) \leq \xi(A, C) + 1$.

Proof. (4) is proved as follows, others can be verified easily.

It is easy to verify $(a \to b) \wedge (b \to a) + (b \to c) \wedge (c \to b) \leq (a \to c) \wedge (c \to a) + 1$ for all $a, b, c \in [0, 1]$. Suppose that A, B and C consist of same atomic formulas $p_1, p_2, \cdots, p_m$. Then

$$\begin{aligned}\xi(A,B)+\xi(B,C) &= \int_{[0,1]^m}\overline{(A\to B)\wedge(B\to A)}(x_1,\cdots,x_m)dx_1\cdots dx_m\\ &\quad+\int_{[0,1]^m}\overline{(B\to C)\wedge(C\to B)}(x_1,\cdots,x_m)dx_1\cdots dx_m\\ &= \int_{[0,1]^m}[(\overline{A}\to\overline{B})\wedge(\overline{B}\to\overline{A})\\ &\quad+(\overline{B}\to\overline{C})\wedge(\overline{C}\to\overline{B})]dx_1\cdots dx_m\\ &\le \int_{[0,1]^m}[(\overline{A}\to\overline{C})\wedge(\overline{C}\to\overline{A})+1]dx_1\cdots dx_m\\ &= \int_{[0,1]^m}\overline{(A\to C)\wedge(C\to A)}(x_1,\cdots,x_m)dx_1\cdots dx_m+1\\ &= \xi(A,C)+1.\end{aligned}$$

5 Pseudo-metric between Formulas in Gödel Logic System

Define a function $\rho : F(S)\times F(S)\to[0,1]$ as follows

$$\rho(A,B)=1-\xi(A,B). \tag{5}$$

Then by Theorem 3, we have following properties.

(1) $\rho(A,A)=0$.

(2) $\rho(A,B)=\rho(B,A)$.

(3) $$\begin{aligned}\rho(A,B)+\rho(B,C) &= 1-\xi(A,B)+1-\xi(B,C)\\ &= 1-(\xi(A,B)+\xi(B,C)-1)\\ &\ge 1-\xi(A,C)\\ &= \rho(A,C).\end{aligned}$$

Therefore, $\rho : F(S)\times F(S)\to[0,1]$ is a pseudo-metric on $F(S)$.

Theorem 4. *Suppose that $A,B\in F(S)$ in Gödel logic system, then we have the following properties.*

(1) *If $A\approx B$, then $\rho(A,B)=0$.*

(2) *If one of A and B is a tautology and other is a contradiction, then $\rho(A,B)=1$.*

(3) *If A is a tautology, then $\rho(A,B)=1-\tau(B)$.*

(4) *$\rho(A,\overline{0})\ge\tau(A)$, where $\overline{0}$ denotes any contradiction.*

Proof. (4) is proved as follows, others can be verified easily.

$$\begin{aligned}\rho(A,\overline{0}) &= 1-\xi(A,\overline{0})\\ &= 1-\tau((A\to\overline{0})\wedge(\overline{0}\to A))\\ &= 1-\int_{[0,1]^m}\overline{(A\to\overline{0})\wedge(\overline{0}\to A)}(x_1,\cdots,x_m)dx_1\cdots dx_m\\ &= 1-\int_{[0,1]^m}(\overline{A}(x_1,\cdots,x_m)\to 0)\wedge(0\to\overline{A}(x_1,\cdots,x_m))dx_1\cdots dx_m\\ &= 1-\int_{[0,1]^m}\overline{A}(x_1,\cdots,x_m)\to 0dx_1\cdots dx_m\\ &= 1-\int_{[0,1]^m}\overline{A\to\overline{0}}(x_1,\cdots,x_m)dx_1\cdots dx_m\\ &= 1-\tau(A\to\overline{0}).\end{aligned}$$

By Corollary 1, we have

$$\tau(A)+\tau(A\to\overline{0})\le\tau(\overline{0})+1=1.$$

i.e.,

$$1 - \tau(A \to \overline{0}) \geq \tau(A).$$

Therefore,$\rho(A, \overline{0}) \geq \tau(A)$.

Remark 2. The equal relation in Theorem 4(4) are not necessary to hold. The counter-example is given as follows.

Consider the formula $A = (p_1 \to p_2) \vee p_1$. It is easy to obtain the function $\overline{A}$ induced by A and $\overline{A \to \overline{0}}$ induced by $A \to \overline{0}$ bellow.

$$\overline{A}(x_1, x_2) = \begin{cases} x_1, & x_1 > x_2 \\ 1, & x_1 \leq x_2 \end{cases}$$

Considering the fact $\overline{A}(x_1, x_2) > 0$ for all $x_1, x_2 \in [0, 1]$, we have

$$\overline{A \to \overline{0}}(x_1, x_2) = \overline{A}(x_1, x_2) \to 0 = 0.$$

Obviously, $\tau(A) = \int_0^1 \int_0^1 \overline{A}(x_1, x_2) dx_1 dx_2 < 1$, while $\rho(A, \overline{0}) = 1 - \tau((A \to \overline{0}) \wedge (\overline{0} \to A)) = 1 - \tau(A \to \overline{0}) = 1 - \int_0^1 \int_0^1 \overline{A \to \overline{0}}(x_1, x_2) dx_1 dx_2 = 1$. i.e., $\rho(A, \overline{0}) \neq \tau(A)$.

References

1. Wang, G.J., Fu, L., Song, J.S.: Theory of truth degrees of propositions in two-valued logic. Science In China, Series A 45(9), 1106–1116 (2002)
2. Wang, G.J.: Introduction to mathematical logic and resolution principle. Science Press, Beijing (2006)
3. Zhang, D.X., Li, L.F.: Syntactic graded method of two-Valued propositional logic formulas. Acta Electronica Sinica 36(2), 325–330 (2008)
4. Li, B.J., Wang, G.J.: Theory of truth degrees of formulas in Łukasiwicz n-valued propositional logic and limit theorem. Science In China, Series F 48(6), 727–736 (2005)
5. Li, J., Li, S.P., Xia, Y.F.: Theory of truth degrees of propositions in n-valued Łukasiwicz logic. Acta Mathematica Sinica 47(4), 769–780 (2004)
6. Ru, Y.M., Wang, G.J.: Distribution of propositional truth degree in some 3-value propositional logic systems. Journal of Shaanxi Normal University 33(4), 6–10 (2005)
7. Wang, G.J., Qin, X.Y., Zhou, X.N.: Theory of quasi-truth degrees of formulas in two-valued predicate logic. Journal of Shaanxi Normal University 33(1), 1–6 (2005)
8. Wang, G.J., Leung, Y.: Integrated semantic and logic metric spaces. Fuzzy Sets and Systems 136, 71–91 (2003)
9. Wang, G.J., Duan, Q.L.: Theory of (n) truth degrees of formulas in modal logic and a consistency theorem. Science In China, Series F 52(1), 70–83 (2009)
10. Wang, G.J., Zhang, W.X.: Consistency degrees of finite theories in Łukasiwicz propositional fuzzy logic. Fuzzy Sets and Systems 149, 275–284 (2005)
11. Zhou, X.N., Wang, G.J.: Consistency degrees of theories in some systems of propositional fuzzy logic. Fuzzy Sets and Systems 152, 321–331 (2005)

12. Zhou, H.J., Wang, G.J.: A new theory consistency index based on deduction theorems in several logic system. Fuzzy Sets and Systems 157, 427–443 (2006)
13. Wu, H.B., Wen, Q.M.: Calculus similar degree and pseudometric in calculus semantics. Journal of Shaanxi Normal University 28(3), 15–19 (2000)
14. Wang, G.J.: Quantified logic (I). Journal of Engineering Mathematics 23(2), 191–215 (2006) (in Chinese)
15. Li, J., Wang, G.J.: Theory of α-truth degree in Gödel n-valued propositional logic. Journal of Software 18(1), 33–39 (2007)
16. Wu, H.B., Yan, M.F.: Generalized semantic HS-rule and MP-rule in Gödel logic system. Journal of Engineering Mathematics 20(3), 56–62 (2003)
17. McNaughton, R.: A theorem about infinite-valued sentential logic. Journal of Symbolic Logic 16, 1–13 (1951)

Finite Topological Models of Modal Logic Systems S4 and S5

Gao-lin Li[1,2] and Luo-shan Xu[1,⋆]

[1] Department of Mathematics, Yangzhou University, Yangzhou 225002, China
e-mail: luoshanxu@hotmail.com

[2] Department of Mathematics, Yancheng Teachers College,
Yancheng 224002, China

Abstract. In modal logic systems **S4** and **S5**, some semantics equivalent theorems on various classes of models are proved. Main results are: (1) A formula is S4 (S5)-theorem if and only if it is globally true in any finite models of S4 (S5); (2) A formula is S4-theorem if and only if it is globally true in all finite topological models; (3) As a corollary of (2), Question 9.1.66 in [4] is affirmatively answered; (4) A formula is S5-theorem if and only if it is globally true in the class FCOT of finite topological models with open sets being closed.

Keywords: Modal logic, model, topological model, filtration, globally true.

1 Introduction

Like classical propositional logic [1], the study of modal logic [2, 3, 4] can be divided into two aspects of syntax and semantics. From the aspect of syntax, modal logic is just a logic system added some modal connectives to classic propositional logic. Different modal logic systems can be obtained by adding some appropriate modal connectives, axioms and rules of inference. However, it becomes more complicated when we compare classical propositional logic and modal logic from semantics. The key point is that interpretations (assignments) of propositional logic involve only one possible world (that is, actual world), while since modal logic includes modal connectives, interpretations (assignments) of modal logic involve a lot of possible worlds. So, even in some fairly simple modal logic systems such as basic modal logic system **K** and its extensions **S4** and **S5**, there are also some basic problems about their semantics to be solved, for example [4, Question 9.1.66].

It is well known that in semantics of modal logic, one can define different concepts of models in terms of different mathematical structures. Typical models are defined by relational structures [2, 3, 4]. A kind of topological

⋆ Corresponding author.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 311–319.
springerlink.com

models given by Mckinsey and tarski in [5] using topological structures. Practically, different classes of models have their advantages and disadvantages. Obviously, finite relational models and some special finite topological models are most convenient ones. So in this paper we have the purpose to make more linkages of relational models and some special topological models, and meanwhile give equivalences of some classes of (finite) models. As a corollary, Question 9.1.66 in [4] mentioned above is affirmatively answered.

2 Preliminaries

For a set W, we use $\wp(W)$ to denote the power set of W. For the non-explicitly stated notions please refer to [2, 6, 7].

A binary relation R on a set W is called
(i) *reflexive* if xRx for all $x \in W$;
(ii) *symmetric* if xRy implies yRx for all $x, y \in W$;
(iii) *transitive* if xRy and yRz implies xRz for all $x, y, z \in W$.
A preorder on W is a relation which is both reflexive and transitive. An equivalence relation is a reflexive, symmetric and transitive relation. If R is an equivalent relation, we use $[x]_R$ to denote an equivalent class of R containing x.

Definition 2.1 [2]**.** *The basic modal language is defined using a set of proposition letters (or proposition symbols or proposition variables) Φ whose elements are usually denoted p, q, r, and so on, and a unary modal operator $\Diamond$ (diamond). The well formed formulas φ of basic modal language are given by the rule*

$$\varphi ::= p \mid \neg\varphi \mid \bot \mid \varphi_1 \ \vee \varphi_2 \mid \Diamond\varphi,$$

where p ranges over elements of Φ. This definition means that a formula is either a proposition letter, the propositional constant falsum (bottom), a negated formula, a disjunction of formulas, or a formula prefixed by a diamond. The set of all formulas is denoted by $\mathrm{Form}(\Diamond, \Phi)$.

Definition 2.2 [2]**.** *A model for the basic modal language is a triple $M = (W, R, V)$, where W is a non-empty set, R is a binary relation on W and V is a function assigning to each proposition letter p in Φ a subset $V(p)$ of W. The function V is called a valuation.*

Definition 2.3 [2, 4]**.** *Suppose w is a state in a model $M = (W, R, V)$. Then we inductively define the notion of a formula φ being satisfied (or true) in M at state w as follows:*
(i) $M, w \models p$ iff $w \in V(p)$, *where* $p \in \Phi$,
(ii) $M, w \models \bot$ *never*,
(iii) $M, w \models \neg\varphi$ iff *not* $M, w \models \varphi$,
(iv) $M, w \models \varphi \vee \psi$ iff $M, w \models \varphi$ or $M, w \models \psi$,
(v) $M, w \models \Diamond\varphi$ iff *for some* $u \in W$ *with* wRu *we have* $M, u \models \varphi$.

A formula φ is *globally* or *universally true* in a model M (notation: $M \models \varphi$) if it is satisfied at all states in M (that is, if $M, w \models \varphi$, for all $w \in W$). A formula φ is *valid* (notation: $\models \varphi$) if it is globally true for all models M.

Definition 2.4 [2]**.** *A set of formulas Σ is* closed under subformulas *(or* subformula closed*) if for all formulas $\varphi, \psi \in Form(\Diamond, \Phi)$,*
(i) $\varphi \vee \psi \in \Sigma \Rightarrow \varphi \in \Sigma$ and $\psi \in \Sigma$,
(ii) $\neg\varphi \in \Sigma \Rightarrow \varphi \in \Sigma$,
(iii) $\Diamond\varphi \in \Sigma \Rightarrow \varphi \in \Sigma$.

Definition 2.5 [3]**.** *A set of formulas $\Sigma(\varphi) \subseteq Form(\Diamond, \Phi)$ is called a subformula set generated by φ, if it satisfies the following four conditions:*
(i) if $\varphi \in \Phi$, then $\Sigma(\varphi) = \{\varphi\}$,
(ii) *if* $\varphi = \psi_1 \vee \psi_2$*, then* $\Sigma(\varphi) = \{\psi_1 \vee \psi_2\} \cup \Sigma(\psi_1) \cup \Sigma(\psi_2)$,
(iii) *if* $\varphi = \neg\psi$*, then* $\Sigma(\varphi) = \{\neg\psi\} \cup \Sigma(\psi)$,
(iv) *if* $\varphi = \Diamond\psi$*, then* $\Sigma(\varphi) = \{\Diamond\psi\} \cup \Sigma(\psi)$.

Remark 2.1. *It is easy to show that $\Sigma(\varphi)$ is subformula closed and finite for all $\varphi \in Form(\Diamond, \Phi)$.*

Definition 2.6 [2, 3]**.** *Let $M = (W, R, V)$ be a model and Σ a subformula closed set of formulas. Let $\equiv_\Sigma$ be the relation on the states of M defined by:*

$$w \equiv_\Sigma v \ \text{ iff } \ for\ all\ \varphi \in \Sigma,\ (M, w \models \varphi \text{ iff } M, v \models \varphi).$$

Note that $\equiv_\Sigma$ is an equivalence relation. We denote the equivalence class of a state w of M with respect to $\equiv_\Sigma$ by $|w|_\Sigma$, or simply by $|w|$ if no confusion will arise.

Let $W_\Sigma = \{|w|_\Sigma : w \in W\}$. Suppose M^f_Σ is any model (W^f, R^f, V^f) such that:
(i) $W^f = W_\Sigma$,
(ii) If wRv then $|w|R^f|v|$,
(iii) If $|w|R^f|v|$ then for all $\Diamond\varphi \in \Sigma$, if $M, v \models \varphi$ then $M, w \models \Diamond\varphi$,
(iv) $V^f(p) = \{|w| : M, w \models p\}$, for all proposition letters p in Σ,
then M^f_Σ is called a *filtration of M through Σ.*

Proposition 2.1 [2]**.** *Let Σ be a finite subformula closed set of basic modal formulas. For any model M, if M^f is a filtration of M through a subformula closed set Σ, then M^f contains at most $2^{Card\Sigma}$ states.*

Theorem 2.1 [2, 3]**.** *Consider the basic modal language. Let $M^f = (W_\Sigma, R^f, V^f)$ be a filtration of M through a subformula close set Σ. Then for all formulas $\varphi \in \Sigma$, and all states w in M, we have that $M, w \models \varphi$ iff $M^f, |w| \models \varphi$.*

3 Finite Models of Systems S4 and S5

It is well known that different modal logic systems can be obtained by taking appropriate restrictions on binary relations R in models $M = (W, R, V)$ of basic modal language. If the relations R are chosen to be preorders, then one gets the model class of the modal logic system **S4**. This model class is also written as S4. Similarly, if the relations R are chosen to be equivalent relations, then one gets the model class S5 of the modal logic system **S5**. Precisely,

$$S4 = \{M = (W, R, V) : R \text{ is a preorder on } W\},$$

$$S5 = \{M = (W, R, V) : R \text{ is an equivalent relation on } W\}.$$

Definition 3.1 [4]**.** *A formula φ is* S4-valid *(resp.,* S5-valid*), written $\models_{S4} \varphi$ (resp., $\models_{S5} \varphi$), if it is globally true in the sense of Definition 2.3 for all models $M \in S4$ (resp., $M \in S5$).*

By the completeness of **S4** and **S5** (i.e., $\vdash_{S4} \varphi \Leftrightarrow \models_{S4} \varphi$ and $\vdash_{S5} \varphi \Leftrightarrow \models_{S5} \varphi$) established in [4, Theorem 9.1.48 and Theorem 9.1.71], we know that the set of all S4-valid formulas is equal to the set of all S4-theorems and that the set of all S5-valid formulas is equal to the set of all S5-theorems. The following two theorems show that when we choose the model classes of finite models for **S4** and **S5**, we also have similar equivalences.

Theorem 3.1. *Let $\varphi \in Form(\Diamond, \Phi)$. Then*

$$\models_{FS4} \varphi \Leftrightarrow \ \models_{S4} \varphi, \tag{1}$$

where $FS4 = \{M = (W, R, V) \in S4 : W \text{ is a finite set}\}$.

Proof. $\Leftarrow$: Trivial.

$\Rightarrow$: For each model $M \in S4$, construct a model $M^f = (W_{\Sigma(\varphi)}, R^t, V^f)$, where $W_{\Sigma(\varphi)}$ and V^f are defined by Definition 2.6 and R^t is defined as follows: for all $|w|, |v| \in W_{\Sigma(\varphi)}$,

$$|w|R^t|v| \Leftrightarrow \ \forall \Diamond\psi \in \Sigma(\varphi), \text{if } M, v \models \psi \vee \Diamond\psi, \text{then } M, w \models \Diamond\psi. \tag{2}$$

Firstly, we show that R^t satisfies Definition 2.6 (ii) (iii) and thus M^f is indeed a filtration of M through $\Sigma(\varphi)$.

To verify Definition 2.6 (ii), suppose wRv. For all $\Diamond\psi \in \Sigma(\varphi)$, if $M, v \models \psi \vee \Diamond\psi$, then $M, v \models \psi$ or $M, v \models \Diamond\psi$. When $M, v \models \psi$, we have $M, w \models \Diamond\psi$ by wRv. When $v \models \Diamond\psi$, there exists v' such that vRv' and $M, v' \models \psi$. Since R is transitive, we have wRv' and $M, w \models \Diamond\psi$. So, $|w|R^t|v|$ and Definition 2.6 (ii) holds.

To verify Definition 2.6 (iii), suppose $|w|R^t|v|$. For all $\Diamond\psi \in \Sigma(\varphi)$, if $M, v \models \psi$, then $v \models \psi \vee \Diamond\psi$ and $w \models \Diamond\psi$ by (2). So, Definition 2.6 (iii) holds.

Secondly, we show that R^t is reflexive and transitive. In fact, for all $|w| \in W_{\Sigma(\varphi)}$, since R is reflexive, we have $|w|R^t|w|$. This means that R^t is reflexive. To show the transitivity of R^t, suppose $|w|R^t|u|$ and $|u|R^t|v|$. For all $\Diamond\psi \in \Sigma(\varphi)$, if $M, v \models \psi \vee \Diamond\psi$, then $M, u \models \Diamond\psi$ and $M, w \models \Diamond\psi$ by (2). So we have $|w|R^t|v|$ by (2) again. This means that R^t is also transitive.

Now, we can conclude that $M^f \in FS4$ by Remark 2.1 and Proposition 2.1. Then by Theorem 2.1, we have $M \models \varphi \Leftrightarrow M^f \models \varphi$. Since $M \in S4$ is arbitrary, we have $\models_{FS4} \varphi \Rightarrow \models_{S4} \varphi$, as desired.

Theorem 3.2. *Let $\varphi \in Form(\Diamond, \Phi)$. Then*

$$\models_{FS5} \varphi \Leftrightarrow \ \models_{S5} \varphi, \tag{3}$$

where $FS5 = \{M = (W, R, V) \in S5 : W \text{ is a finite set}\}$.

Proof. $\Leftarrow$: Trivial.

$\Rightarrow$: For each model $M \in S5$, construct a model $M^f = (W_{\Sigma(\varphi)}, R^e, V^f)$, where $W_{\Sigma(\varphi)}$ and V^f are defined by Definition 2.6 and R^e is defined as follows: for all $|w|, |v| \in W_{\Sigma(\varphi)}$, $|w|R^e|v|$ iff the following two conditions are satisfied:

$$\forall \Diamond\psi \in \Sigma(\varphi) : \ (M, v \models \psi \vee \Diamond\psi) \Rightarrow (M, w \models \Diamond\psi), \tag{4}$$

$$\forall \Diamond\psi \in \Sigma(\varphi) : \ (M, w \models \Diamond\psi) \Rightarrow (M, v \models \Diamond\psi). \tag{5}$$

Firstly, we show that R^e satisfies Definition 2.6 (ii) (iii) and thus M^f is indeed a filtration of M through $\Sigma(\varphi)$.

To verify Definition 2.6 (ii), suppose wRv. On the one hand, for all $\Diamond\psi \in \Sigma(\varphi)$, if $M, v \models \psi \vee \Diamond\psi$, then $M, v \models \psi$ or $M, v \models \Diamond\psi$. When $M, v \models \psi$, we have $M, w \models \Diamond\psi$ by wRv. When $M, v \models \Diamond\psi$, there exists v' such that vRv' and $M, v' \models \psi$. Since R is transitive, we have wRv', and then $M, w \models \Diamond\psi$. On the other hand, for all $\Diamond\psi \in \Sigma(\varphi)$, if $M, w \models \Diamond\psi$, then there exists w' such that wRw' and $M, w' \models \psi$. By the symmetry of R, we have vRw, and then by the transitivity of R, we have vRw', hence $M, v \models \Diamond\psi$. So, $|w|R^e|v|$ and Definition 2.6 (ii) is verified.

To verify Definition 2.6 (iii), suppose $|w|R^t|v|$. For all $\Diamond\psi \in \Sigma(\varphi)$, if $M, v \models \psi$, then $M, v \models \psi \vee \Diamond\psi$, and then $M, w \models \Diamond\psi$ by (4). Thus, Definition 2.6 (iii) is verified.

Secondly, we show that R^e is reflexive, transitive and symmetric. In fact, for all $|w| \in W_{\Sigma(\varphi)}$, since R is reflexive, we have $|w|R^e|w|$. This means that R^e is reflexive. To show the transitivity of R^e, suppose $|w|R^e|u|$ and $|u|R^e|v|$. On the one hand, for all $\Diamond\psi \in \Sigma(\varphi)$, if $M, v \models \psi \vee \Diamond\psi$, then $M, u \models \Diamond\psi$, and $M, w \models \Diamond\psi$ by (4). On the other hand, for all $\Diamond\psi \in \Sigma(\varphi)$, if $w \models \Diamond\psi$, then $u \models \Diamond\psi$, and $v \models \Diamond\psi$ by (5). So we have $|w|R^e|v|$ by the definition of R^e. This means that R^e is transitive. To show the symmetry of R^e, suppose $|w|R^e|v|$. For all $\Diamond\psi \in \Sigma(\varphi)$. On the one hand, if $M, w \models \psi \vee \Diamond\psi$, then $M, w \models \psi$ or $M, w \models \Diamond\psi$, and then $M, w \models \Diamond\psi$. So we have that $M, v \models \Diamond\psi$ by (5). On the other hand, if $M, v \models \Diamond\psi$, then $M, v \models \psi \vee \Diamond\psi$, and then $M, w \models \Diamond\psi$. So we have $|v|R^e|w|$.

Now, we can conclude that $M^f \in FS5$ by Remark 2.1 and Proposition 2.1. Then by Theorem 2.1, we have $M \models \varphi \Leftrightarrow M^f \models \varphi$. Since $M \in S5$ is arbitrary, we have $\models_{FS5} \varphi \Rightarrow \models_{S5} \varphi$, as desired.

4 Finite Topological Models of System S4 and S5

Definition 4.1. *A topological model for the basic model language is a triple $M = (W, \mathcal{T}, V)$, where W is a non-empty set, $\mathcal{T}$ is a topology on W and V is a function assigning to each proposition letter p in Φ a subset $V(p)$ of W. The function V is called a valuation.*

It is easy to see that topological models are special minimal models in the sense of [3, Definition 7.1] defined by special mappings.

Definition 4.2. *Suppose w is a state in a topological model $M = (W, \mathcal{T}, V)$. Then we inductively define the notion of a formula φ being* satisfied *(or* true*) in M at state w as follows:*

(i) $M, w \models_T p$ iff $w \in V(p),\ p \in \Phi$.
(ii) $M, w \models_T \perp$ *never.*
(iii) $M, w \models_T \neg\varphi$ iff *not* $M, w \models_T \varphi$.
(iv) $M, w \models_T \varphi \vee \psi$ iff $M, w \models_T \varphi$ or $M, w \models_T \psi$.
(v) $M, w \models_T \Diamond\varphi$ iff *for any open neighborhood U of w, there exists $u \in U$ such that $M, u \models_T \varphi$.*

A formula φ is *globally* or *universally true* in a topological model M (notation: $M \models_T \varphi$) if it is satisfied at all states in M (that is, if $M, w \models_T \varphi$ for all $w \in W$). A formula φ is *topologically valid* (notation: $\models_T \varphi$) if it is globally true for all topological models M.

Remark 4.1. *It is easy to verify that the concepts of globally true and topologically valid formulas in Definition 4.2 are equivalent to the corresponding concepts in [4, Definition 9.1.49]. Moreover, by [4, Th.9.1.53 and Th.9.1.58], the topological soundness and topological completeness of system* **S4***, we can conclude that a formula is S4-theorem if and only if this formula is S4-valid if and only if the formula is topologically valid, namely,*

$$\vdash_{S4} \varphi \Leftrightarrow \ \models_{S4} \varphi \Leftrightarrow \ \models_T \varphi, \quad \varphi \in Form(\Diamond, \Phi). \tag{6}$$

Let TM be the class of all topological models and $M = (W, \mathcal{T}, V) \in TM$. Then M is called an Alexandrov topological model if $\mathcal{T}$ is an Alexandrov topology. The class of all Alexandrov topological models is denoted by AT. If W is a finite set, then $M = (W, \mathcal{T}, V)$ is called a finite topological model. The class of all finite topological models is denoted by FT. A formula $\varphi \in$Form$(\Diamond, \Phi)$ is called Alexandrov topologically valid (notation: $\models_{AT} \varphi$) if $M \models_T \varphi$ for all $M \in AT$. Similarly, we can define the concept of finite topologically valid formulas (notation: $\models_{FT} \varphi$).

It is well know that for any finite topological space $(W, \mathcal{T})$, $\mathcal{T}$ is closed under arbitrary intersections and is an Alexandrov topology. So we have

$$\models_T \varphi \Rightarrow \ \models_{AT} \varphi \Rightarrow \ \models_{FT} \varphi, \quad \varphi \in \mathrm{Form}(\Diamond, \Phi). \tag{7}$$

Lemma 4.1. *Let W be a non-empty set, R a binary relation on W.*

(i) *If R is a preorder, then $\alpha(R) = \{A \subseteq W : \forall x, y \in W, xRy \wedge x \in A \Rightarrow y \in A\}$ is an Alexandrov topology on W.*

(ii) *If R is an equivalence relation, then $\mathcal{T}(R) = \{A \subseteq W : [x]_R \subseteq A$, for all x in $A\}$ is a topology on W with every open set being also closed.*

(iii) *If $\mathcal{T}$ is a topology on W, then $R(\mathcal{T}) = \{(x, y) \in W \times W : \mathrm{Cl}(\{x\}) = \mathrm{Cl}(\{y\})\}$ is an equivalence relation.*

Proof. (i) By [4, Remark 9.1.62].

(ii) It is easy to verify $\mathcal{T}(R)$ is a topology on W. Moreover if $A \in \mathcal{T}(R)$ and $x \in \mathrm{Cl}(A)$, then we have $[x]_R \cap A \neq \emptyset$. Then there is a $y \in W$ such that xRy and $y \in A$. Thus $x \in [y]_R \subseteq A$, hence $A = \mathrm{Cl}(A)$. So every open set is also closed with respect to $\mathcal{T}(R)$.

(iii) Directed verification.

Theorem 4.1. *Let $\varphi \in Form(\Diamond, \Phi)$. Then φ is an $S4-$theorem if and only if φ is finite topologically valid, namely,*

$$\vdash_{S4} \varphi \Leftrightarrow \ \models_{FT} \varphi, \quad \forall\ \varphi \in Form(\Diamond, \Phi). \tag{8}$$

Proof. $\Rightarrow$: By (6) and (7).

$\Leftarrow$: By (1), it suffices to show that

$$\models_{FT} \varphi \Rightarrow \models_{FS4} \varphi, \quad \forall\ \varphi \in \mathrm{Form}(\Diamond, \Phi). \tag{9}$$

For each model $M = (W, R, V) \in FS4$, we can get a topological model $M^\alpha = (W, \alpha(R), V)$, where $\alpha(R)$ is defined as in Lemma 4.1(i). It is clear that $M^\alpha = (W, \alpha(R), V) \in FT$.

We claim that for all $w \in W$ and $\varphi \in \mathrm{Form}(\Diamond, \Phi)$,

$$M, w \models \varphi \Leftrightarrow M^\alpha, w \models_T \varphi. \tag{10}$$

We prove this claim inductively with respect to the complexity of φ. If $\varphi = p \in \Phi$, then

$$M, w \models p \Leftrightarrow w \in V(p) \Leftrightarrow M^\alpha, w \models_T p. \tag{11}$$

Suppose that (10) is true for formulas with the number of connectives not more than k. Then, for formula φ with $k+1$ connectives, we prove that (10) is also true. We divide the proof into the following three cases: (a) $\varphi = \neg\psi$, (b) $\varphi = \psi_1 \vee \psi_2$ and (c) $\varphi = \Diamond\psi$.

For case (a), we have

$M, w \models \neg\psi \Leftrightarrow$ not $M, w \models \psi$
$\Leftrightarrow$ not $M^\alpha, w \models_T \psi$ (by the assumption of induction)
$\Leftrightarrow M^\alpha, w \models_T \neg\psi$.

For case (b), we have

$M, w \models \psi_1 \vee \psi_2 \Leftrightarrow M, w \models \psi_1$ or $M, w \models \psi_2$
$\Leftrightarrow M^\alpha, w \models_T \psi_1$ or $M^\alpha, w \models_T \psi_2$ (by the assumption of induction)
$\Leftrightarrow M^\alpha, w \models_T \psi_1 \vee \psi_2$.

For case (c), we have

$M, w \models \Diamond\psi \Leftrightarrow$ there is $u \in W$ such that wRu and $M, u \models \psi$

$\Leftrightarrow$ for all open neighborhood U of w in $(W, \alpha(R))$, there is $u \in U$ such that $M^\alpha, u \models_T \psi$ (by the assumption of induction)

$\Leftrightarrow M^\alpha, w \models_T \Diamond\psi$.

So, by principle of induction, (10) is true for all $w \in W$ and $\varphi \in$Form$(\Diamond, \Phi)$. Then by the arbitrariness of $M \in FS4$, we conclude that (9) is true.

Theorem 4.2. *For all $\varphi \in$ Form$(\Diamond, \Phi)$, we have*

$$\vdash_{S4} \varphi \Leftrightarrow \ \models_{S4} \varphi \Leftrightarrow \ \models_T \varphi \Leftrightarrow \ \models_{AT} \varphi \Leftrightarrow \ \models_{FT} \varphi. \tag{12}$$

Proof. It follows immediately from (6), (7) and (8).

Remark 4.2. *Question 9.1.66 in* [4] now is affirmatively answered by Theorem 4.2.

Next we pass to consider finite models in system **S5**. Set $FCOT = \{(W, \mathcal{T}, V) \in FT : A \text{ is closed for all } A \in \mathcal{T}\}$. For every $M = (W, \mathcal{T}, V) \in FCOT$, we can get a model $M^R = (W, R(\mathcal{T}), V) \in FS5$, where $R(\mathcal{T})$ is defined as in Lemma 4.1(iii).

Theorem 4.3. *Let $\varphi \in$Form$(\Diamond, \Phi)$. Then φ is a $S5-$theorem if and only if φ is globally true in all topological models $M \in FCOT$, namely,*

$$\vdash_{S5} \varphi \Leftrightarrow \models_{FCOT} \varphi, \quad \forall\, \varphi \in \textit{Form}(\Diamond, \Phi).$$

Proof. $\Rightarrow$: By Theorem 3.2 and [4, Theorem 9.1.71], the completeness of system **S5**, it suffices to show that $\models_{FS5} \varphi \Rightarrow \models_{FCOT} \varphi$.

If $M = (W, \mathcal{T}, V) \in FCOT$, then the model $M^R = (W, R(\mathcal{T}), V) \in FS5$.

We claim that for all $w \in W$ and $\varphi \in$ Form$(\Diamond, \Phi)$

$$M, w \models_T \varphi \Leftrightarrow M^R, w \models \varphi. \tag{13}$$

We prove this claim inductively with respect to the complexity of φ. If $\varphi = p \in \Phi$, then

$$M, w \models_T p \Leftrightarrow w \in V(p) \Leftrightarrow M^R, w \models p.$$

Suppose that (13) is true for formulas with the number of connectives not more than k. Then, for formula φ with $k+1$ connectives, we prove that (13) is also true. We divide the proof into the following three cases: (a) $\varphi = \neg\psi$, (b) $\varphi = \psi_1 \vee \psi_2$ and (c) $\varphi = \Diamond\psi$. For case (a) and case (b), the proves are similar to the corresponding cases in the proof of Theorem 4.1 and omitted.

For case (c), we have

$M, w \models_T \Diamond\psi \Leftrightarrow$ for all open neighborhood U of w in $(W, \mathcal{T})$, there is $u \in U$ such that $M, u \models_T \psi$

$\Leftrightarrow$ for the smallest open neighborhood V of w in$(W, \mathcal{T})$, there is $v \in V$ such that $M, v \models_T \psi$

$\Leftrightarrow$ there is $v \in W$ such that $wR(\mathcal{T})v$ and$M^R, v \models \psi$ (by the assumption of induction)

$\Leftrightarrow M^R, w \models \Diamond\psi$.

So (13) is true for all $w \in W$ and $\varphi \in$Form$(\Diamond, \Phi)$. Thus $\models_{FS5} \varphi \Rightarrow \models_{FCOT} \varphi$.

$\Leftarrow$: By Theorem 3.2, it suffices to prove that $\models_{FCOT} \varphi \Rightarrow \models_{FS5} \varphi$.

For each model $M = (W, R, V) \in FS5$, since R is an equivalent relation on W, we see that $M^{\mathcal{T}} = (W, \mathcal{T}(R), V) \in FCOT$, where $\mathcal{T}(R)$ is defined as in Lemma 4.1(ii).

It is a routine work to show by induction with respect to the complexity of φ that $M^{\mathcal{T}}, w \models_T \varphi \Leftrightarrow M, w \models \varphi$. So $\models_{FCOT} \varphi \Rightarrow \models_{FS5} \varphi$, as desired.

Acknowledgements. This work is supported by the NSF of China (10371106, 60774073), of Jiangsu University Education Section (08KJD110008) and of Yancheng Teachers College (07YCKL070).

References

1. Wang, G., Zhou, H.: Introduction to mathematical logic and resolution principle, 2nd edn. Science Press, Beijing (2009)
2. Blackburn, P., Rijke, M., Venema, Y.: Modal logic. Cambridge University Press, Cambridge (2001)
3. Chellas, B.F.: Modal logic: An introduction. Cambridge University Press, New York (1980)
4. Wang, G.: Non-classical mathematical logic and approximate reasoning, 2nd edn. Science Press, Beijing (2008)
5. Mckinsey, J., Tarski, A.: The algebra of topology. Annals of Mathematics 45, 141–191 (1944)
6. Gierz, G., et al.: Continuous lattices and domains. Cambridge University Press, New York (2003)
7. Engelking, R.: General topology. Polish Scientific Publishers, Warszaw (1977)

Researches on (n) Truth Degrees in Modal Logic

Mei-hua Cui[1,2] and Luo-shan Xu[2,⋆]

[1] Department of Mathematics, Yancheng Teachers College, Yancheng 224002, China
[2] Department of Mathematics, Yangzhou University, Yangzhou 225002, China
e-mail: luoshanxu@hotmail.com

Abstract. More equalities and inequalities about (n) truth degrees are deduced in basic modal logic systems K. Some properties about the pseudo distance are given. More theorems about (n) truth degrees and the pseudo distance in temporal logic are proved. Simple proves of continuities of modal logic operators in (n) modal logic metric spaces are given. We also studied divergence in modal logic metric spaces and proved the equivalence of three given approximate reasoning theories.

Keywords: Modal logic metric spaces, temporal logic, truth degree, pseudo distance, divergence degree.

1 Introduction

Rosser and Turquetter [1] in 1952 posed the idea of distinguishing credible degrees of logic formulas. To complete this idea, Pavelka posed a logic theory to stratify fully in the frame of lattice valued propositional logic [2]. Since 90's of last century, various stratified logic concepts have been studied and developed and many results have been obtained [3]-[10]. Based on mean probability, Wang in [4] firstly posed the concept of truth degrees of logic formulas and the theory of logic metric spaces in classical logic systems. And the theory of quantitative logic [9]-[10] formed step by step. The establishing of quantitative logic makes it come true to stratify logic concepts. Quantitative logic also combines symbolic logic and numerical computing, as well as approximate computing in computational mathematics together, broadening application areas of formal logic. Recently, Wang and Duan in [11] posed the concepts of (n) truth degrees in basic modal logic system K in terms of finite semantical models. They also defined modal logic metric spaces and gave some theory of approximate reasoning. Based on [11], we in this paper will deduce more equalities and inequalities about (n) truth degrees and give

⋆ Corresponding author.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 321–330.
springerlink.com

more properties of pseudo metric in basic modal logic systems K. Some new results about (n) truth degrees and pseudo distance in temporal logic will also be obtained. Simple proves of continuities of modal logic operators in the (n) modal logic metric space are given. We also study divergence in modal logic metric spaces and prove the equivalence of three given approximate reasoning theories.

2 Preliminaries

Modal logic [12] is an extension of classical logic by adding some modal operators and axioms related to modal operators. Semantics of basic modal logic are given in terms of Kripke semantical models [13]. The non-explicitly stated notions and symbols in this paper, please refer to [11, 15].

Definition 2.1 [14]**.** *Let $\diamond$ be a unary modal operator, Φ a set of atomic propositions. $\diamond$-type modal formulas are constructed as following:*

$$\varphi := p \mid \perp \mid \neg\varphi \mid \varphi_1 \vee \varphi_2 \mid \diamond\varphi,\ p \in \Phi,$$

i.e., (i) $\forall p \in \Phi, p$ is a $\diamond$-type modal formulas;

(ii) $\perp$ is a $\diamond$-type modal formulas ($\perp$ represents the refusion formulas);

(iii) If $\varphi, \varphi_1, \varphi_2$ are $\diamond$-type modal formulas, then $\neg\varphi, \diamond\varphi$ and $\varphi_1 \vee \varphi_2$ are also $\diamond$-type modal formulas.

We use Form $(\diamond, \Phi)$ to denote the set of $\diamond$-type modal formulas based on Φ and $\Box$ to denote the dual modal operator of $\diamond$, $\varphi \wedge \psi$ denotes $\neg(\neg\varphi \vee \neg\psi)$, and $\varphi \to \psi$ denotes $\neg\varphi \vee \psi$, and $\Box\varphi = \neg\diamond\neg\varphi$.

Definition 2.2 [14]**.** *Modal logic system K is consisting of*

(i) *Formulas: all the $\diamond$-type modal formulas in Form$(\diamond, \Phi)$.*

(ii) *Axioms:*

1° *All the tautologies in two-valued propositional logic based on Φ and $\neg\perp$;*

2° *distributive axioms: $\Box(p \to q) \to (\Box p \to \Box q),\ p, q \in \Phi$.*

(iii) *Inference Rules*:

1° *MP-rules: from φ and $\varphi \to \psi$ one can deduce ψ;*

2° *Generalization Rules: from φ one can deduce $\Box\varphi$;*

3° *Uniformly substitutions: namely, if $\varphi = f(p_1 \cdots p_n)$ and $h : \Phi \to$Form$(\diamond, \Phi)$ is a map, then from φ one can deduce $f(h(p_1) \cdots h(p_n)), (p_1, \cdots, p_n \in \Phi)$.*

It follows from distributive axioms and uniformly substitutions that $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$.

If φ is a theorem in K, then we write that $\vdash_K \varphi$, or briefly $\vdash \varphi$. When $\vdash \varphi \to \psi$ and $\vdash \psi \to \varphi$ hold, we say that φ and ψ are provable equivalent, written $\varphi \sim \psi$.

Definition 2.3 [15]**.** *Let $\Gamma \subseteq Form(\diamond, \Phi)$, $\varphi \in Form(\diamond, \Phi)$. Then Γ is called a theory. If Γ is non-empty and there are $\psi_1, \cdots, \psi_n \in \Gamma$ such that $\vdash \psi_1 \wedge \cdots \wedge \psi_n \to \varphi$, then we say that φ can be deduced by Γ, or φ is a Γ-conclusion, written $\Gamma \vdash \varphi$. We use $D(\Gamma)$ to denote the set of all Γ-conclusion.*

Definition 2.4 [14]**.** *A model for basic modal logic system K is a triple $M = (W, R, V)$, where W is a non-empty set, R is a binary relation on W (i.e., $R \subseteq W \times W$), and $V : \Phi \to \wp(W)$ (the power set of W) is a map called an evaluation.*

Definition 2.5 [14]**.** *Let $M = (W, R, V)$ be a model for basic modal logic system K, $w \in W$ and $\varphi \in Form(\diamond, \Phi)$. We define inductively that φ is true at the point w of the model M or w satisfies φ, written $M, w \models \varphi$, as follows:*
(i) $M, w \models p$ iff $w \in V(p), P \in \Phi$,
(ii) $M, w \models \perp$ *never holds,*
(iii) $M, w \models \neg\varphi$ iff $M, w \models \varphi$ *not hold,*
(iv) $M, w \models \varphi \vee \psi$ iff $M, w \models \varphi$ *or* $M, w \models \psi$,
(v) $M, w \models \diamond\varphi$ iff there are $u \in W$ *such that* $R_{w,u}$ *and* $M, u \models \varphi$, *where* $R_{w,u}$ *means that* $(w, u) \in R$.

If $\forall w \in W$, $M, w \models \varphi$ holds, then we say that φ is globally true in M, written $M \models \varphi$. If for all M of models for K, $M \models \varphi$ holds, then we say φ is a valid formula, written $\models \varphi$. When $\models \varphi \to \psi$ and $\models \psi \to \varphi$ hold, we say φ and ψ are logic equivalent, written $\varphi \approx \psi$.

Let $\varphi \in$Form$(\diamond, \Phi)$. Then the construction of φ evolves only finite atomic propositions. So, we assume that Φ is a finite set and consider only finite models for K in the sequel. For a finite model $M = (W, R, V)$, if $|W| = n$, then we take $W_n = \{1, 2, \cdots, n\}$ as a standard domain, and the set of all the finite models (W_n, R, V) based on W_n is also a finite one, and written briefly $\mathcal{M}_n$ (note: it is related to Φ). And we write $\mathcal{M}_f = \cup_{n=1}^{\infty} \mathcal{M}_n$.

Definition 2.6 [11]**.** *Let $\varphi \in Form(\diamond, \Phi)$.*
(i) *Let $M = (W, R, V) \in \mathcal{M}_f$. Define the truth degree $\tau_M(\varphi)$ of φ with respect to M as follows: $\tau_M(\varphi) = \mu(\{w \in W \mid M, w \models \varphi\})$ or $\tau_M(\varphi) = \mu(V(\varphi)), (V(\varphi) = \{w \in W \mid M, w \models \varphi\})$, where μ is the mean probabilistic measure on finite set W.*
(ii) *The (n) truth degree $\tau_n(\varphi)$ of φ is defined as follows:*

$$\tau_n(\varphi) = \frac{1}{|\mathcal{M}_n|} \sum \{\tau_M(\varphi) \mid M \in \mathcal{M}_n\}.$$

Remark 2.1 *.(The invariant property of the (n) truth degree) Above summation is a finite one and exists. In addition, if one adds some atomic propositions to Φ, then the (n) truth degree of the same formula is unchanged. Precisely, if $\varphi \in Form(\diamond, \Phi)$ and $\Phi_0 \supseteq \Phi$ is finite, then*

$$\tau_n(\varphi) = \frac{1}{|\mathcal{M}_n(\Phi_0)|} \sum \{\tau_M(\varphi) \mid M \in \mathcal{M}_n(\Phi_0)\}.$$

This means that Definition 2.6 is meaningful. So, if necessary or for sake of conveniences, one can add/deduct some finite atomic propositions to Φ to guarantee meaningfulness of matters in question and leave the (n) truth degree of the same formula unchanged.

Definition 2.7 [11]**.** *Let $M = (W, R, V)$, $W = \{1, 2, \cdots, n\}$, $\varphi, \psi \in Form(\diamond, \Phi)$. Set*

$$\xi_n(\varphi, \psi) = \tau_n((\varphi \to \psi) \wedge (\psi \to \varphi)).$$

We say that $\xi_n(\varphi, \psi)$ is the (n) similarity of φ and ψ.

Lemma 2.1 [11]**.** *Let $\varphi, \psi \in Form(\diamond, \Phi)$. Then*
(i) $V(\neg\varphi) = W - V(\varphi)$. (ii) $V(\varphi \vee \psi) = V(\varphi) \cup V(\psi)$,
(iii) $V(\varphi \wedge \psi) = V(\varphi) \cap V(\psi)$, (iv) $V(\varphi \to \psi) = (W - V(\varphi)) \cup V(\psi)$,
(v) $V(\diamond\varphi) = \{w \in W \mid R[w] \cap V(\varphi) \neq \emptyset\}$, *where* $R[w] = \{u \in W \mid R_{w,u}\}$.

Lemma 2.2 [11]**.** *Let $\varphi, \psi \in Form(\diamond, \Phi)$. Then $\forall\ n = 1, 2, \cdots$,*
(i) $0 \leq \tau_n(\varphi) \leq 1$, (ii) $\tau_n(\varphi) + \tau_n(\neg\varphi) = 1$,
(iii) *If* $\models \varphi$, *then* $\tau_n(\varphi) = 1$, (iv) *If* $\varphi \sim \psi$, *then* $\tau_n(\varphi) = \tau_n(\psi)$.

Lemma 2.3 [11]**.** *Let $\varphi, \psi, \varphi_1, \varphi_2, \varphi_3 \in Form(\diamond, \Phi)$, $\alpha, \beta \in [0, 1]$. Then*
(i) $\tau_n(\varphi \vee \psi) + \tau_n(\varphi \wedge \psi) = \tau_n(\varphi) + \tau_n(\psi)$,
(ii) *If* $\tau_n(\varphi) \geq \alpha, \tau_n(\varphi \to \psi) \geq \beta$, *then* $\tau_n(\psi) \geq \alpha + \beta - 1$,
(iii) *If* $\tau_n(\varphi_1 \to \varphi_2) \geq \alpha, \tau_n(\varphi_2 \to \varphi_3) \geq \beta$, *then* $\tau_n(\varphi_1 \to \varphi_3) \geq \alpha + \beta - 1$.

Lemma 2.4 [11]**.** *Let $\varphi, \psi, \varphi_1, \varphi_2, \varphi_3 \in Form(\diamond, \Phi)$. Then*
(i) *if* $\varphi \sim \psi$, *then* $\xi_n(\varphi, \psi) = 1$,
(i) $\xi_n(\varphi_1, \varphi_2) + \xi_n(\varphi_2, \varphi_3) \leq 1 + \xi_n(\varphi_1, \varphi_3)$.

Lemma 2.5 [11]**.** *Let $\varphi, \psi \in Form(\diamond, \Phi)$, define ρ_n: $Form(\diamond, \Phi)^2 \to [0, 1]$, such that.*

$$\rho_n(\varphi, \psi) = 1 - \xi_n(\varphi, \psi).$$

Then ρ_n is a pseudo distance on $Form(\diamond, \Phi)$, and $(Form(\diamond, \Phi), \rho_n)$ is called an (n) modal logic metric space.

Temporal logic is a special modal logic and modal operator $\diamond$ has the following meaning, $\diamond p$ means that p holds in some occasions of future, and $\Box p$ means that p holds in every occasion of future. For a finite model $M = (W, R, V)$ of temporal logic with $W = \{1, 2, \cdots, n\}$, binary relation R is a chain. So, we can choose $R = \{(w, u) \mid w, u \in W_n = \{1, 2, \cdots, n\}, w \leq u\}$. We use $\mathcal{M}_{n,R}$ to denote the set of such temporal models. Define the (n) truth degree $\tau_{n,R}(\varphi)$ of φ w. r. t. $\mathcal{M}_{n,R}$ such that

$$\tau_{n,R}(\varphi) = \frac{1}{|\mathcal{M}_{n,R}|} \sum \{\tau_M(\varphi) \mid M \in \mathcal{M}_{n,R}\}.$$

Lemma 2.6 [11]**.** *Let $p \in \Phi$. Then, for all $n = 1, 2, \cdots$ we have*
(i) $\tau_{n,R}(\Box p) = \frac{2}{n} - \frac{1}{n2^{n-1}}$,
(ii) $\tau_{n,R}(\diamond p) = 1 - \frac{2}{n} + \frac{1}{n2^{n-1}}$,
(iii) $\tau_{n,R}(\diamond p \wedge p) = \frac{1}{2} - \frac{1}{n} + \frac{1}{n2^{n}}$,
(iv) $\tau_{n,R}(\diamond p \wedge \Box p) = \frac{1}{n} - \frac{1}{n2^{n-1}}$.

3 The (n) Truth Degree of a Modal Formula

Proposition 3.1. *Let $\varphi, \psi, \varphi_1, \varphi_2, \varphi_3 \in$Form$(\diamond, \Phi)$. Then $\forall\ n = 1, 2, \cdots$,*
(i) $\tau_n(\varphi \to \psi) = \tau_n(\varphi \wedge \psi) - \tau_n(\varphi) + 1$,
(ii) $\tau_n(\varphi_1 \to (\varphi_2 \vee \varphi_3)) + \tau_n(\varphi_1 \to (\varphi_2 \wedge \varphi_3)) = \tau_n(\varphi_1 \to \varphi_2) + \tau_n(\varphi_1 \to \varphi_3)$,
(iii) $\tau_n(\varphi_1 \vee \varphi_2 \to \varphi_3) + \tau_n(\varphi_1 \wedge \varphi_2 \to \varphi_3) = \tau_n(\varphi_1 \to \varphi_3) + \tau_n(\varphi_2 \to \varphi_3)$,
(iv) $\tau_n(\varphi_1 \vee \varphi_2 \to \varphi_1 \wedge \varphi_2) = \tau_n(\varphi_1 \to \varphi_2) + \tau_n(\varphi_2 \to \varphi_1) - 1$.

Proof. (i) By Lemma 2.3(i), we have that $\tau_n(\varphi \to \psi) + \tau_n(\varphi) = \tau_n(\neg\varphi \vee \psi) + \tau_n(\varphi) = \tau_n(\neg\varphi \vee \psi \vee \varphi) + \tau_n((\neg\varphi \vee \psi) \wedge \varphi) = 1 + \tau_n(\varphi \wedge \psi)$. So, $\tau_n(\varphi \to \psi) = \tau_n(\varphi \wedge \psi) - \tau_n(\varphi) + 1$.

(ii) It follows from (i) that

$$\begin{aligned}
&\tau_n(\varphi_1 \to \varphi_2 \vee \varphi_3) + \tau_n(\varphi_1 \to \varphi_2 \wedge \varphi_3)\\
&\quad = \tau_n(\varphi_1 \wedge (\varphi_2 \vee \varphi_3)) - \tau_n(\varphi_1) + 1 + \tau_n(\varphi_1 \wedge \varphi_2 \wedge \varphi_3) - \tau_n(\varphi_1) + 1\\
&\quad = \tau_n((\varphi_1 \wedge \varphi_2) \vee (\varphi_1 \wedge \varphi_3)) + \tau_n((\varphi_1 \wedge \varphi_2) \wedge (\varphi_1 \wedge \varphi_3)) - 2\tau_n(\varphi_1) + 2\\
&\quad = \tau_n(\varphi_1 \wedge \varphi_2) + \tau_n(\varphi_1 \wedge \varphi_3) - 2\tau_n(\varphi_1) + 2\\
&\quad = \tau_n(\varphi_1 \to \varphi_2) + \tau_n(\varphi_1 \to \varphi_3).
\end{aligned}$$

The proof of (iii) is similar to that of (ii).

(iv)

$$\begin{aligned}
&\tau_n(\varphi_1 \vee \varphi_2 \to \varphi_1 \wedge \varphi_2)\\
&\quad = \tau_n((\varphi_1 \vee \varphi_2) \wedge (\varphi_1 \wedge \varphi_2)) - \tau_n(\varphi_1 \vee \varphi_2) + 1\\
&\quad = 2\tau_n(\varphi_1 \wedge \varphi_2) - \tau_n(\varphi_1) - (\varphi_2) + 1\\
&\quad = \tau_n(\varphi_1 \to \varphi_2) + \tau_n(\varphi_2 \to \varphi_1) - 1.
\end{aligned}$$

By 3.1 (i), we can pithily prove Lemmas 2.3 (ii), (iii) and 2.4 (ii).

A proof of Lemma 2.3 (ii): $\tau_n(\psi) \geq \tau_n(\varphi \wedge \psi) = \tau_n(\varphi \to \psi) + \tau_n(\varphi) - 1$.

A proof of Lemma 2.3 (iii):

$$\begin{aligned}
&\tau_n(\varphi_1 \to \varphi_2) + \tau_n(\varphi_2 \to \varphi_3) - 1\\
&\quad = \tau_n(\varphi_1 \wedge \varphi_2) + \tau_n(\varphi_2 \wedge \varphi_3) - \tau_n(\varphi_1) - \tau_n(\varphi_2) + 1\\
&\quad = \tau_n((\varphi_1 \vee \varphi_3) \wedge \varphi_2) + \tau_n(\varphi_1 \wedge \varphi_2 \wedge \varphi_3) - \tau_n(\varphi_1) - \tau_n(\varphi_2) + 1\\
&\quad \leq \tau_n(\varphi_1 \wedge \varphi_2 \wedge \varphi_3) - \tau_n(\varphi_1) + 1\\
&\quad \leq \tau_n(\varphi_1 \wedge \varphi_3) - \tau_n(\varphi_1) + 1\\
&\quad = \tau_n(\varphi_1 \to \varphi_3).
\end{aligned}$$

A proof of Lemma 2.4(ii): Since $\tau_n((\varphi_1 \to \varphi_2) \vee (\varphi_2 \to \varphi_1)) = \tau_n((\neg\varphi_1 \vee \varphi_2) \vee (\neg\varphi_2 \vee \varphi_1)) = 1$, we have that

$$\begin{aligned}
&\xi_n(\varphi_1, \varphi_2) + \xi_n(\varphi_2, \varphi_3)\\
&\quad = \tau_n((\varphi_1 \to \varphi_2) \wedge (\varphi_2 \to \varphi_1)) + \tau_n((\varphi_2 \to \varphi_3) \wedge (\varphi_3 \to \varphi_2))\\
&\quad = \tau_n(\varphi_1 \to \varphi_2) + \tau_n(\varphi_2 \to \varphi_1) - 1 + \tau_n(\varphi_2 \to \varphi_3) + \tau_n(\varphi_3 \to \varphi_2) - 1\\
&\quad = [\tau_n(\varphi_1 \to \varphi_2) + \tau_n(\varphi_2 \to \varphi_3) - 1] + [\tau_n(\varphi_3 \to \varphi_2) + \tau_n(\varphi_2 \to \varphi_1) - 1].
\end{aligned}$$

By Lemma 2.3 (iii) we have that

$$\xi_n(\varphi_1, \varphi_2) + \xi_n(\varphi_2, \varphi_3) \leq \tau_n(\varphi_1 \to \varphi_3) + \tau_n(\varphi_3 \to \varphi_1) = 1 + \xi_n(\varphi_1, \varphi_3).$$

Proposition 3.2. *Let* $\varphi_1, \varphi_2, \varphi_3 \in Form(\diamond, \Phi)$, $\alpha, \beta \in [0,1]$. *Then* $\forall\ n = 1, 2, \cdots$,

(i) *if* $\tau_n(\varphi_1 \to \varphi_2) \geq \alpha, \tau_n(\varphi_1 \to \varphi_3) \geq \beta$, *then* $\tau_n(\varphi_1 \to \varphi_2 \wedge \varphi_3) \geq \alpha + \beta - 1, \tau_n(\varphi_1 \to \varphi_2 \vee \varphi_3) \geq \alpha + \beta - 1$;

(ii) *if* $\tau_n(\varphi_1 \to \varphi_3) \geq \alpha, \tau_n(\varphi_2 \to \varphi_3) \geq \beta$, *then* $\tau_n(\varphi_1 \wedge \varphi_2 \to \varphi_3) \geq \alpha + \beta - 1, \tau_n(\varphi_1 \vee \varphi_2 \to \varphi_3) \geq \alpha + \beta - 1$.

Proof. (i) $\tau_n(\varphi_1 \to \varphi_2 \wedge \varphi_3) = \tau_n(\neg\varphi_1 \vee (\varphi_2 \wedge \varphi_3)) = \tau_n(\neg\varphi_1 \vee \varphi_2) + \tau_n(\neg\varphi_1 \vee \varphi_3) - \tau_n(\neg\varphi_1 \vee \varphi_2) \wedge (\neg\varphi_1 \vee \varphi_3) \geq \tau_n(\varphi_1 \to \varphi_2) + \tau_n(\varphi_1 \to \varphi_3) - 1 = \alpha + \beta - 1$.

Similar arguments can prove that $\tau_n(\varphi_1 \to \varphi_2 \vee \varphi_3) \geq \alpha + \beta - 1$ and (ii).

Proposition 3.3. *Let* $p \in \Phi$. *Then in temporal logic, for all* $n = 1, 2, \cdots$ *we have*

$$\tau_{n,R}(\Box p \wedge p) = \frac{3}{2n} - \frac{1}{n2^{n-1}}.$$

Proof. Since the formula only evolves an atomic proposition p, we can take $\Phi = \{p\}$, and it is easy to see that the cardinality $|\mathcal{M}_{n,R}| = 2^n$. Thus, $V(\Box p \wedge p) = V(\Box p) \cap V(p) = \{w \in W_n \mid \forall u \in W_n$ with $w < u$ one has $u \in V(P)\} \cap V(P)$.

(i) If $V(\Box p) = \{n\}$, then $V(P)$ does not contain n and $V(\Box p \wedge p) = V(\Box p) \cap V(p) = \emptyset$.

(ii) If $V(\Box p) = \uparrow k$ $(2 \leq k \leq n-1)$, then $V(p)$ has 2^{k-1} possible choices and max$V(p) = n$. So,

$V(\Box p \wedge p) = V(\Box p) \cap V(p) = \{n, n-1, \cdots, k\}$ and $\mu(V(\Box p \wedge p)) = \frac{n-k+1}{n}$.

(iii) If $V(\Box p) = W$, then $V(p)$ has only two choices and max$V(p) = n$. So $\mu(V(\Box p \wedge p)) = 1$ and $\tau_{n,R}(\Box p \wedge p) = \frac{1}{|\mathcal{M}_{n,R}|} \sum\{\tau_M(\Box p \wedge p) \mid M \in \mathcal{M}_{n,R}\} = \frac{1}{2^n}(\sum_{k=2}^{n-1} \frac{n-k+1}{n} 2^{k-1} + 2) = \frac{3}{2n} - \frac{1}{n2^{n-1}}$.

Proposition 3.4. *Let* $p \in \Phi$. *Then in temporal logic, for all* $n = 1, 2, \cdots$,

(i) $\tau_{n,R}(p \to \Box p) = \frac{1}{2} + \frac{3}{2n} - \frac{1}{n2^{n-1}}$,
(ii) $\tau_{n,R}(\Box p \to p) = 1 - \frac{1}{2n}$,
(iii) $\tau_{n,R}(p \to \diamond p) = 1 - \frac{1}{n} + \frac{1}{n2^n}$,
(iv) $\tau_{n,R}(\diamond p \to p) = \frac{1}{2} + \frac{1}{n} - \frac{1}{n2^n}$,
(v) $\tau_{n,R}(\Box p \to \diamond p) = 1 - \frac{1}{n}$,
(vi)$\tau_{n,R}(\diamond p \to \Box p) = \frac{3}{n} - \frac{1}{n2^{n-2}}$.

Proof. (i) By Lemma 2.5 and Propositions 3.1(i)and 3.3 we have that

$$\tau_{n,R}(p \to \Box p) = \tau_{n,R}(p \wedge \Box p) - \tau_{n,R}(p) + 1 = \frac{3}{2n} - \frac{1}{n2^{n-1}} - \frac{1}{2} + 1 = \frac{1}{2} + \frac{3}{2n} - \frac{1}{n2^{n-1}}.$$

Similar argument can prove (ii)-(vi).

4 Metric Spaces of (n) Modal Logic

Proposition 4.1. *Let $\varphi_1, \varphi_2, \in Form(\diamond, \Phi)$. Then $\forall\ n = 1, 2, \cdots$ we have*

(i) $\rho_n(\varphi_1, \varphi_2) = 2 - \tau_n(\varphi_1 \to \varphi_2) - \tau_n(\varphi_2 \to \varphi_1) = \tau_n(\varphi_1 \vee \varphi_2) - \tau_n(\varphi_1 \wedge \varphi_2)$,

(ii) $\rho_n(\neg\varphi_1, \neg\varphi_2) = \rho_n(\varphi_1 \vee \varphi_2, \varphi_1 \wedge \varphi_2) = \rho_n(\varphi_1 \to \varphi_2, \varphi_2 \to \varphi_1) = \rho_n(\varphi_1, \varphi_2)$.

Proof. (i) $\rho_n(\varphi_1, \varphi_2) = 1 - \tau_n((\varphi_1 \to \varphi_2) \wedge (\varphi_2 \to \varphi_1)) = 1 - \tau_n(\varphi_1 \to \varphi_2) - \tau_n(\varphi_2 \to \varphi_1) + \tau_n((\varphi_1 \to \varphi_2) \vee (\varphi_2 \to \varphi_1)) = 2 - \tau_n(\varphi_1 \to \varphi_2) - \tau_n(\varphi_2 \to \varphi_1)$. By Prop 3.1(i),

$$\rho_n(\varphi_1, \varphi_2) = 2 - 2\tau_n(\varphi_1 \wedge \varphi_2) + \tau_n(\varphi_1) + \tau_n(\varphi_2) - 2 = \tau_n(\varphi_1 \vee \varphi_2) - \tau_n(\varphi_1 \wedge \varphi_2).$$

(ii) $\rho_n(\neg\varphi_1, \neg\varphi_2) = 2 - \tau_n(\neg\varphi_1 \to \neg\varphi_2) - \tau_n(\neg\varphi_2 \to \neg\varphi_1) = 2 - \tau_n(\varphi_2 \to \varphi_1) - \tau_n(\varphi_1 \to \varphi_2) = \rho_n(\varphi_1, \varphi_2)$.

$\rho_n(\varphi_1 \vee \varphi_2, \varphi_1 \wedge \varphi_2) = \tau_n((\varphi_1 \vee \varphi_2) \vee (\varphi_1 \wedge \varphi_2)) - \tau_n((\varphi_1 \vee \varphi_2) \wedge (\varphi_1 \wedge \varphi_2)) = \tau_n(\varphi_1 \vee \varphi_2) - \tau_n(\varphi_1 \wedge \varphi_2) = \rho_n(\varphi_1, \varphi_2)$.

$\rho_n(\varphi_1 \to \varphi_2, \varphi_1 \to \varphi_2) = \tau_n((\varphi_1 \to \varphi_2) \vee (\varphi_2 \to \varphi_1)) - \tau_n((\varphi_1 \to \varphi_2) \wedge (\varphi_2 \to \varphi_1)) = 1 - \tau_n((\varphi_1 \to \varphi_2) \wedge (\varphi_2 \to \varphi_1)) = \rho_n(\varphi_1, \varphi_2)$.

By Propositions 4.1(ii) and 3.4, one has the following corollary immdiately.

Corollary 4.1. *If $p \in \Phi$, then in temporal logic, for all $n = 1, 2, \cdots$ we have that*

(i) $\rho_{n,R}(p \to \Box p, \Box p \to p) = \rho_{n,R}(p \vee \Box p, p \wedge \Box p) = \rho_{n,R}(\neg p, \neg\Box p)$
$= \rho_{n,R}(p, \Box p) = \frac{1}{2} - \frac{1}{n} + \frac{1}{n2^{n-1}}$,

(ii) $\rho_{n,R}(p \to \diamond p, \diamond p \to p) = \rho_{n,R}(p \vee \diamond p, p \wedge \diamond p) = \rho_{n,R}(\neg p, \neg \diamond p)$
$= \rho_{n,R}(p, \diamond p) = \frac{1}{2}$,

(iii) $\rho_{n,R}(\diamond p \to \Box p, \Box p \to \diamond p) = \rho_{n,R}(\diamond p \vee \Box p, \diamond p \wedge \Box p) = \rho_{n,R}(\neg \diamond p, \neg\Box p)$
$= \rho_{n,R}(\diamond p, \Box p) = 1 - \frac{2}{n} + \frac{1}{n2^{n-2}}$.

Proposition 4.2. *Let $\varphi_1, \varphi_2, \in Form(\diamond, \Phi)$. Then $\forall\ n = 1, 2, \cdots$ we have*

(i) $\rho_n(\varphi_1 \wedge \varphi_2, \varphi_1) = 1 - \tau_n(\varphi_1 \to \varphi_2)$,
(ii) $\rho_n(\varphi_1 \vee \varphi_2, \varphi_1) = 1 - \tau_n(\varphi_2 \to \varphi_1)$,
(iii) $\rho_n(\varphi_1 \to \varphi_2, \varphi_1) = 1 - \tau_n(\varphi_1 \wedge \varphi_2)$,
(iv) $\rho_n(\varphi_1 \to \varphi_2, \varphi_2) = 1 - \tau_n(\varphi_1 \vee \varphi_2)$,
(v) $\rho_n(\varphi_1 \to \varphi_2, \varphi_1 \vee \varphi_2) = 1 - \tau_n(\varphi_2)$,
(vi) $\rho_n(\varphi_1 \to \varphi_2, \varphi_1 \wedge \varphi_2) = 1 - \tau_n(\varphi_1)$.

Proof. We prove (i), (iii) and (v) only. The other items can be similarly proved.

(i) $\rho_n(\varphi_1 \wedge \varphi_2, \varphi_1) = 2 - \tau_n(\varphi_1 \wedge \varphi_2 \to \varphi_1) - \tau_n(\varphi_1 \to \varphi_1 \wedge \varphi_2) = 1 - \tau_n(\varphi_1 \to \varphi_2)$.

(iii) $\rho_n(\varphi_1 \to \varphi_2, \varphi_1) = \tau_n((\varphi_1 \to \varphi_2) \vee \varphi_1) - \tau_n((\varphi_1 \to \varphi_2) \wedge \varphi_1) = \tau_n(\neg\varphi_1 \vee \varphi_2 \vee \varphi_1) - \tau_n((\neg\varphi_1 \vee \varphi_2) \wedge \varphi_1) = 1 - \tau_n(\varphi_1 \wedge \varphi_2)$.

(v) $\rho_n(\varphi_1 \to \varphi_2, \varphi_1 \vee \varphi_2) = \tau_n((\varphi_1 \to \varphi_2) \vee (\varphi_1 \vee \varphi_2)) - \tau_n((\varphi_1 \to \varphi_2) \wedge (\varphi_1 \vee \varphi_2)) = \tau_n(\neg\varphi_1 \vee \varphi_2 \vee \varphi_1 \vee \varphi_2) - \tau_n((\neg\varphi_1 \vee \varphi_2) \wedge (\varphi_1 \vee \varphi_2)) = 1 - \tau_n(\varphi_2)$.

Proposition 4.3. *Let* $\varphi_1, \varphi_2, \varphi_3 \in$*Form*$(\diamond, \Phi)$. *Then*
(i) $\rho_n(\varphi_1 \vee \varphi_3, \varphi_2 \vee \varphi_3) \leq \rho_n(\varphi_1, \varphi_2)$,
(ii) $\rho_n(\varphi_1 \wedge \varphi_3, \varphi_2 \wedge \varphi_3) \leq \rho_n(\varphi_1, \varphi_2)$,
(iii) $\rho_n(\varphi_1 \to \varphi_3, \varphi_2 \to \varphi_3) \leq \rho_n(\varphi_1, \varphi_2)$,
(iv) $\rho_n(\varphi_3 \to \varphi_1, \varphi_3 \to \varphi_2) \leq \rho_n(\varphi_1, \varphi_2)$.

Proof. (i) $\rho_n(\varphi_1 \vee \varphi_3, \varphi_2 \vee \varphi_3)$
$= 2 - \tau_n((\varphi_1 \vee \varphi_3) \to (\varphi_2 \vee \varphi_3)) - \tau_n((\varphi_2 \vee \varphi_3) \to (\varphi_1 \vee \varphi_3))$
$= 2 - 2\tau_n((\varphi_1 \vee \varphi_3) \wedge (\varphi_2 \vee \varphi_3)) + \tau_n(\varphi_1 \vee \varphi_3) + \tau_n(\varphi_2 \vee \varphi_3) - 2$
$= -2\tau_n((\varphi_1 \wedge \varphi_2) \vee \varphi_3) + \tau_n(\varphi_1) + \tau_n(\varphi_2) + 2\tau_n(\varphi_3) - \tau_n(\varphi_1 \wedge \varphi_3) - \tau_n(\varphi_2 \wedge \varphi_3)$
$= \tau_n(\varphi_1) + \tau_n(\varphi_2) - 2\tau_n(\varphi_1 \wedge \varphi_2) - \tau_n(\varphi_1 \wedge \varphi_3) - \tau_n(\varphi_2 \wedge \varphi_3) + 2\tau_n(\varphi_1 \wedge \varphi_2 \wedge \varphi_3)$
$\leq \tau_n(\varphi_1) + \tau_n(\varphi_2) - 2\tau_n(\varphi_1 \wedge \varphi_2)$
$= \tau_n(\varphi_1 \vee \varphi_2) - \tau_n(\varphi_1 \wedge \varphi_2)$
$= \rho_n(\varphi_1, \varphi_2)$.
Similar arguments can prove(ii) and (iii).
(iv) $\rho_n(\varphi_3 \to \varphi_1, \varphi_3 \to \varphi_2) = 2 - \tau_n((\varphi_3 \to \varphi_1) \to (\varphi_3 \to \varphi_2)) - \tau_n((\varphi_3 \to \varphi_2) \to (\varphi_3 \to \varphi_1))$.
Note that $\tau_n((\varphi_1 \to \varphi_2) \to ((\varphi_3 \to \varphi_1) \to (\varphi_3 \to \varphi_2))) = 1$. By Lemma 2.3 (ii), we have
$\tau_n((\varphi_3 \to \varphi_1) \to (\varphi_3 \to \varphi_2)) \geq \tau_n((\varphi_1 \to \varphi_2)$,
$\tau_n((\varphi_3 \to \varphi_2) \to (\varphi_3 \to \varphi_1)) \geq \tau_n((\varphi_2 \to \varphi_1)$,
So, $\rho_n(\varphi_3 \to \varphi_1, \varphi_3 \to \varphi_2) \leq 2 - \tau_n(\varphi_1 \to \varphi_2) - \tau_n(\varphi_2 \to \varphi_1) = \rho_n(\varphi_1, \varphi_2)$.

Proposition 4.4. *Let* $\varphi_1, \varphi_2, \varphi_3, \varphi_4 \in$*Form*$(\diamond, \Phi)$. *Then* $\forall\ n = 1, 2, \cdots$ *we have*
(i) $\rho_n(\varphi_1 \vee \varphi_2, \varphi_3 \vee \varphi_4) \leq \rho_n(\varphi_1, \varphi_3) + \rho_n(\varphi_2, \varphi_4)$.
(ii) $\rho_n(\varphi_1 \wedge \varphi_2, \varphi_3 \wedge \varphi_4) \leq \rho_n(\varphi_1, \varphi_3) + \rho_n(\varphi_2, \varphi_4)$.
(iii) $\rho_n(\varphi_1 \to \varphi_2, \varphi_3 \to \varphi_4) \leq \rho_n(\varphi_1, \varphi_3) + \rho_n(\varphi_2, \varphi_4)$.

Proof. (i) $\rho_n(\varphi_1 \vee \varphi_2, \varphi_3 \vee \varphi_4) = 1 - \xi_n(\varphi_1 \vee \varphi_2, \varphi_3 \vee \varphi_4)$
$\leq 2 - \xi_n(\varphi_1 \vee \varphi_2, \varphi_2 \vee \varphi_3) - \xi_n(\varphi_2 \vee \varphi_3, \varphi_3 \vee \varphi_4)$
$= \rho_n(\varphi_1 \vee \varphi_2, \varphi_2 \vee \varphi_3) + \rho_n(\varphi_2 \vee \varphi_3, \varphi_3 \vee \varphi_4) \leq \rho_n(\varphi_1, \varphi_3) + \rho_n(\varphi_2, \varphi_4)$.
Similar arguments can proof (ii) and (iii).

By Propositions 4.1(ii) and 4.4, one has immediately the following theorem on the continuities of modal operators in modal logic metric spaces.

Theorem 4.1. *In modal logic metric spaces (Form*$(\diamond, \Phi), \rho_n)$*, unary operator "*$\neg$*" and binary operators "*$\vee$*", "*$\wedge$*" and "*$\to$*" are all continuous w.r.t.* ρ_n.

Definition 4.1. *Let* $\Gamma \subseteq$*Form*$(\diamond, \Phi)$. *Set*

$$\mathrm{div}_n(\Gamma) = \sup\{\rho_n(\varphi, \psi) \mid \varphi, \psi \in D(\Gamma)\}.$$

We call $div_n(\Gamma)$ *the divergent degree of theory* Γ. *When* $div_n(\Gamma) = 1$ *we call* Γ *to be fully divergent.*

Theorem 4.2 *Let* $\Gamma = \{\varphi_1, \cdots, \varphi_m\} \subseteq$*Form*$(\diamond, \Phi)$. *Then*

$$div_n(\Gamma) = 1 - \tau_n(\varphi_1 \wedge \cdots \wedge \varphi_m).$$

Proof. For all $\psi \in D(\Gamma)$, by Definition 2.3, we have $\vdash (\varphi_1 \wedge \cdots \wedge \varphi_m) \to \psi$. It follows from Lemma 2.3 (ii) that

$\tau_n(\psi) \geq \tau_n(\varphi_1 \wedge \cdots \wedge \varphi_m) + \tau_n((\varphi_1 \wedge \cdots \wedge \varphi_m) \to \psi) - 1 = \tau_n(\varphi_1 \wedge \cdots \wedge \varphi_m)$.
By the arbitrariness of ψ we know that among formulas in $D(\Gamma)$, $\varphi_1 \wedge \cdots \wedge \varphi_m \in D(\Gamma)$ has the smallest truth degree.

By Proposition 4.1(i), $\rho_n(\varphi, \psi) = \tau_n(\varphi \vee \psi) - \tau_n(\varphi \wedge \psi)$. Taking φ to be a theorem, we have $\tau_n(\varphi \vee \psi) = 1$, $\tau_n(\varphi \wedge \psi) = \tau_n(\psi)$. So, $\sup\{\rho_n(\varphi, \psi) \mid \varphi, \psi \in D(\Gamma)\} = 1 - \min\{\tau_n(\psi) \mid \psi \in D(\Gamma)\}$, hence $\mathrm{div}_n(\Gamma) = 1 - \tau_n(\varphi_1 \wedge \cdots \wedge \varphi_m)$.

By Theorem 4.2, if $\Gamma =$Form$(\diamond, \Phi)$, then Γ is fully divergent.

Definition 4.2 [11] **.** *Let $\Gamma \subseteq$Form$(\diamond, \Phi)$, $\varphi \in$Form$(\diamond, \Phi)$ and $\varepsilon > 0$.*

(i) *If $\rho_n(\varphi, D(\Gamma)) =inf\{\rho_n(\varphi, \psi) \mid \psi \in D(\Gamma)\} < \varepsilon$, then we call φ a Γ conclusion of type-I by error less than ε, written $\varphi \in D(\Gamma, I, \varepsilon)$.*

(ii) *If 1-sup$\{\tau_n(\psi \to \varphi) | \psi \in D(\Gamma)\} < \varepsilon$, then we call φ a Γ conclusion of type-II by error less than ε, written $\varphi \in D(\Gamma, II, \varepsilon)$.*

(iii) *If $\inf\{H(D(\Gamma), D(\Sigma)) \mid \Sigma \subseteq$ Form$(\diamond, \Phi), \Sigma \vdash \varphi\} < \varepsilon$, where H is Hausdorff distance on (Form$(\diamond, \Phi), \rho_n)$, then we call φ a Γ conclusion of type-III by error less than ε, written $\varphi \in D(\Gamma, III, \varepsilon)$.*

Lemma 4.1 [11] **.** *Let $\Gamma \subseteq$Form$(\diamond, \Phi)$, $\varphi \in$Form$(\diamond, \Phi)$, $\varepsilon > 0$. Then $\varphi \in D(\Gamma, I, \varepsilon)$ iff $\varphi \in D(\Gamma, II, \varepsilon)$.*

Lemma 4.2 [11] **.** *Let $\Gamma \subseteq$Form$(\diamond, \Phi)$, $\varphi \in$Form$(\diamond, \Phi)$ and $\varepsilon > 0$. If $\varphi \in D(\Gamma, III, \varepsilon)$, then $\varphi \in D(\Gamma, I, \varepsilon)$.*

Theorem 4.3. *Let $\Gamma \subseteq$Form$(\diamond, \Phi)$, $\varphi \in$Form$(\diamond, \Phi)$ and $\varepsilon > 0$. If $\varphi \in D(\Gamma, I, \varepsilon)$, then $\varphi \in D(\Gamma, III, \varepsilon)$.*

Proof. If $\varphi \in D(\Gamma, I, \varepsilon)$, then $\rho_n(\varphi, D(\Gamma)) < \varepsilon$. Let $\Sigma' = \Gamma \cup \{\varphi\}$. Then $\Sigma' \subseteq$Form$(\diamond, \Phi)$, $\Sigma' \vdash \varphi$, $D(\Gamma) \subseteq D(\Sigma')$.

Firstly, we prove $H^*(D(\Gamma), D(\Sigma')) < \varepsilon$. For all $\psi \in D(\Gamma)$ we have $\rho_n(\psi, D(\Sigma')) = 0$. So $H^*(D(\Gamma), D(\Sigma')) = \sup\{\rho_n(\psi, D(\Sigma') \mid \psi \in D(\Gamma)\} = 0 < \varepsilon$. Secondly we prove $H^*(D(\Sigma'), D(\Gamma)) < \varepsilon$. For all $\varphi' \in D(\Sigma')$, if $\psi \in D(\Gamma)$, then there is $\{\psi_1, \cdots, \psi_j\} \subseteq \Gamma$ such that $\vdash \psi_1 \wedge \cdots \wedge \psi_j \wedge \varphi \to \varphi'$. Set $\psi^* = \psi \wedge \psi_1 \wedge \cdots \wedge \psi_j$. Then $\vdash \psi^* \to \psi$, $\vdash \psi^* \wedge \varphi \to \varphi'$. Set also $\psi' = \psi^* \vee \varphi'$. Then $\vdash \psi^* \to \psi'$. It follows from $\psi^* \in D(\Gamma)$ and MP-rules that $\psi' \in D(\Gamma)$. By Proposition 4.2(ii) we have

$\rho_n(\varphi', \psi') = \rho_n(\varphi', \psi^* \vee \varphi') = 1 - \tau_n(\psi^* \to \varphi')$,

$1 - \tau_n(\psi^* \to \varphi') \leq 1 - \tau_n(\psi^* \to \psi^* \wedge \varphi) = 1 - \tau_n(\psi^* \to \varphi) \leq 1 - \tau_n(\psi \to \varphi)$,

$\rho_n(\varphi', \psi') \leq 1 - \tau_n(\psi \to \varphi) \leq 1 - \tau_n((\psi \to \varphi) \wedge (\varphi \to \psi)) = \rho_n(\varphi, \psi)$.

Thus $\rho_n(\varphi', D(\Gamma)) \leq \rho_n(\varphi, D(\Gamma))$ and $H^*(D(\Sigma'), D(\Gamma)) = \sup\{\rho_n(\varphi', D(\Gamma) \mid \varphi' \in D(\Sigma')\} \leq \rho_n(\varphi, D(\Gamma)) < \varepsilon$.

To sum up above, we see that $H(D(\Gamma), D(\Sigma')) < \varepsilon$ and

$$\inf\{H(D(\Gamma), D(\Sigma)) \mid \Sigma \subseteq \text{Form}(\diamond, \Phi), \Sigma \vdash \varphi\} \leq H(D(\Gamma), D(\Sigma')) < \varepsilon.$$

So, $\varphi \in D(\Gamma, III, \varepsilon)$.

By Theorem 4.3 and Lemmas 4.1 and 4.2, it is easy to see that the three theories of of approximate reasoning schemes in Definition 4.2 are equivalent.

Acknowledgements. This work is supported by the NSF of China (10371106, 60774073) and the NSF of Jiangsu University Education Section (08KJD110008).

References

1. Rosser, J.B., Turquette, A.R.: Many-valued logics. North-Holland, Amsterdam (1952)
2. Pavelka, J.: On fuzzy logic (I, II, III). Z. Math Logic Grundlagen Math. 25, 45–52, 119–134, 447–464 (1979)
3. Ying, M.: A logic for approximate reasoning. J. Symb. Logic 59, 830–837 (1994)
4. Wang, G., Fu, L., Song, J.: Theory of truth degrees of propositions in two-valued logic. Science in China (Series A: Mathematics) 45(9), 1106–1116 (2002)
5. Wang, G., Leung, Y.: Integrated semantics and logic metric spaces. Fuzzy Sets and Systems 136(1), 71–91 (2003)
6. Wang, G., Zhang, W.: Consistency degrees of finite theories in lukasiewicz propositional fuzzy logic. Fuzzy Sets and Systems 149, 275–284 (2005)
7. Zhou, X., Wang, G.: Consistency degrees of theories in some systems propositional fuzzy logic. Fuzzy Sets and Systems 152, 321–331 (2005)
8. Zhou, H., Wang, G.: A new theory consisteny index based on deduction theorems in several logic systems. Fuzzy Sets and Systems 157, 427–443 (2006)
9. Li, B., Wang, G.: Theory of truth degrees of formulas in Lukasiewicz n-valued propositional logic and a limit theorem. Science in China (Series F: Information Sciences) 48(6), 727–736 (2005)
10. Wang, G.: Introduction to mathematical logic and resolution principle, 2nd edn. Science Press, Beijing (2006)
11. Wang, G., Duan, Q.: Theory of (n) truth degrees of formulas in modal logic and a consistency theorem. Science in China (Series F: Information Sciences) 52(1), 70–83 (2009)
12. Chagrov, Zakharyas, Chev1: Modal logic. MIT Press, Cambridge (1998)
13. Kripke, S.: Semantic analysis of modal logic. Zeitxchrift für Mathematische Logik und Grundlagen der Mathematik 9(1), 67–96 (1963)
14. Blackburn, P., Rijke, M., Venema, Y.: Modal logic. Cambridge University Press, Cambridge (2001)
15. Wang, G.: Non-classical Mathematical Logic and Approximate Reasoning, 2nd edn. Science Press, Beijing (2008)

Lattice-Valued Truth Degree in Łukasiewicz Propositional Fuzzy Logic

Dong-xiao Zhang

School of Information Science and Technology, Xiamen University,
Xiamen 361000, Fujian Province, P.R. China
e-mail: zdx1980@hotmail.com
Department of Mathematics, School of Science, Jimei University,
Xiamen 361021, Fujian Province, P.R. China

Abstract. The study of formula truth degree based on the grading idea has been a hot topic in some common logic systems, such as classical two-valued propositional logic, many-valued propositional logic, predicate logic, fuzzy propositional logic and model logic. So far, almost all definitions of truth degree are given on the unit interval $[0,1]$ whose structures are seldom considered. This paper try to define formula truth degree on MV-algebra, a bounded distributive lattice, as lattice-valued truth degree. Besides profit from plenty inherent properties of MV-algebra and its generalization of unit interval $[0,1]$, lattice-valued truth degree discussed as follows may contribute to establishing truth degree theory about lattice-valued proposition logic.

Keywords: Lattice-valued truth degree, lattice-valued proposition logic, MV-algebra, Łukasiewicz propositional fuzzy logic.

1 Introduction

At present, the study of formula truth degree based on the grading idea has been a hot topic in some common logic systems, such as classical two-valued propositional logic [1-2], many-valued propositional logic [3-5], predicate logic [6], fuzzy propositional logic [7,2], model logic [8]. Moreover, based on the truth degree of formula, many new theories, say consistent degree of theory Γ [9-11], resemblance degree between two formulae [12], are proposed. It can be seen from this that formula truth degree plays an important role in the quantitative logic [13], whose basic concept is very truth degree. Otherwise [14] and [15] try to give a standard of the fine truth degree of formula.

So far, almost all definitions of truth degree are given on the unit interval $[0,1]$ whose structures are seldom considered. This paper try to define formula truth degree on MV-algebra [16], a bounded distributive lattice, as

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 331–337.
springerlink.com

lattice-valued truth degree. Otherwise, [17-18] propose lattice-valued proposition logic base on the lattice implication algebra [19], and [20] proves that lattice implication algebra and MV-algebra are the equivalent algebra system. Therefore, Besides profit from plenty inherent properties of MV-algebra and its generalization of unit interval $[0, 1]$, lattice-valued truth degree discussed as follows can be imitated to establish truth degree theory about lattice-valued proposition logic.

2 Preliminaries

Let $S = \{p_1, p_2, \cdots\}$ be a countable set, and $\neg$ and $\rightarrow$ be unary and binary logic connectives respectively. Write $F(S)$ be the free algebra of type $(\neg, \rightarrow)$ generated by S.Elements of $F(S)$ are called propositions or formulae. In Łukasiewicz propositional fuzzy logic (briefly, **Łuk**), there are four axiom schemes as follows:

(**Lu1**) $A \rightarrow (B \rightarrow A)$.
(**Lu2**) $(A \rightarrow B) \rightarrow ((B \rightarrow C) \rightarrow (A \rightarrow C))$.
(**Lu3**) $((A \rightarrow B) \rightarrow B) \rightarrow ((B \rightarrow A) \rightarrow A)$.
(**Lu4**) $(\neg A \rightarrow \neg B) \rightarrow (B \rightarrow A)$.

The deduction rule in **Łuk** is Modus Ponens (briefly, MP), i.e., B can be deduced from A and $A \rightarrow B$. Suppose that $\Gamma \subset F(S)$ and $A \in F(S)$, then A is a Γ-conclusion if A can be deduced from Γ and preceding axioms by using MP within finite steps, and denoted by $\Gamma \vdash A$. In case $\Gamma = \emptyset$, $\Gamma \vdash A$ can be abbreviated as $\vdash A$ and A is called a theorem of **Łuk**.On the contrary, A is called a refutable formula if $\neg A$ is a theorem. It is said that A and B are provably equivalent and denoted by $A \sim B$ if both $\vdash A \rightarrow B$ and $\vdash B \rightarrow A$ hold. The new connectives $\wedge$, $\vee$, $\oplus$ and $\otimes$ are usually introduced in **Łuk** as follows

$$\begin{aligned} A \vee B = (A \rightarrow B) \rightarrow B,\ A \wedge B = \neg (\neg A \vee \neg B), \\ A \oplus B = \neg A \rightarrow B, A \otimes B = \neg (\neg A \oplus \neg B). \end{aligned} \tag{1}$$

MV-algebra theory [16], which possesses many good properties, is the algebra theory matching with Łukasiewicz logic system. Following is its simplified definition.

Definition 1. *[2]* $(X, \oplus, ', 0)$ *is called MV-algebra, if the following conditions are satisfied:*
(i) $(X, \oplus, 0)$ *is a commutative semigroup.*
(ii) $x \oplus 0' = 0'$.
(iii) $(x')' = x$.
(iv) $(x' \oplus y)' \oplus y = (y' \oplus x)' \oplus x$.

Define relation $\leq$ as $x \leq y$ if and only if $x' \oplus y = 0'$, then $(X, \leq)$ becomes a bounded distributive lattice and

$$x \vee y = (x' \oplus y)' \oplus y, x \wedge y = (x' \vee y')', x, y \in X. \tag{2}$$

Otherwise, define two binary operations $\otimes$ and $\rightarrow$ as follows

$$x \otimes y = (x' \oplus y')', x \rightarrow y = x' \oplus y. \tag{3}$$

Remark 1. The symbol $\rightarrow, \vee, \wedge, \oplus$ and $\otimes$ in equation (1) are the logic connectives in **Łuk**, while the ones in equation (2) and (3) are the binary operators on MV-algebra. It isn't confusing although no distinction are made between them. So no differentiation is required.

3 Lattice-Valued Truth Degree

Following discussions are confined to Łukasiewicz propositional fuzzy logic **Łuk**.

Definition 2. *Let L be a MV-algebra. Then $\tau : F(S) \rightarrow L$ is called a lattice-valued truth degree on L, briefly L-valued truth degree, if the following conditions are satisfied:*
(i) If $\vdash A$, then $\tau(A) = 1, A \in F(S)$.
(ii) $\tau(\neg A) = (\tau(A))', A \in F(S)$.
(iii) $\tau(A \oplus B) \oplus \tau(A \otimes B) = \tau(A) \oplus \tau(B),\ A, B \in F(S)$.

In following, Symbol L denotes some MV-algebra and τ denotes a lattice-valued truth degree on L except for extra illumination.

Considering the fact that $\neg (A \oplus B) \sim \neg A \otimes \neg B$, $\neg (A \otimes B) \sim \neg A \oplus \neg B$, $\forall A, B \in F(S)$, and $(x \oplus y)' = x' \otimes y', (x \otimes y)' = x' \oplus y', \forall x, y \in L$, we have following equivalent characterization of L-valued truth degree.

Theorem 1. *$\tau : F(S) \rightarrow L$ is L-valued truth degree, if and only if the following conditions are satisfied for any $A, B \in F(S)$:*
(i) If $\vdash A$, then $\tau(A) = 1$.
(ii) $\tau(\neg A) = (\tau(A))'$.
(iii) $\tau(A \oplus B) \otimes \tau(A \otimes B) = \tau(A) \otimes \tau(B)$.

There are same truth degree for provably equivalent formulae, following results show that they have same L-valued truth degree.

Theorem 2. *Suppose that $A, B \in F(S)$.*
(i) If $\vdash A \rightarrow B$, then $\tau(A) \leq \tau(B)$.
(ii) If $A \sim B$, then $\tau(A) = \tau(B)$.

Proof. (i) If $\vdash A \rightarrow B$, then $\vdash \neg A \oplus B$,thus $\tau(\neg A \oplus B) = 1$. It is concluded from Definition 2 that

$$\tau(A)' \oplus \tau(B) = \tau(\neg A) \oplus \tau(B) = \tau(\neg A \oplus B) \oplus \tau(\neg A \otimes B) = 1 \oplus \tau(\neg A \otimes B) = 1.$$

Therefore,$\tau(A) \leq \tau(B)$.
(ii) is the result of (i).

Taking advantage of preceding results and Definition 2, it is easy to verify following relations.

Proposition 1. *Following relations hold in* ***Łuk****.*

(i) $\tau(A)\otimes\tau(B)\le\tau(A\otimes B)\le\tau(A\wedge B)$
$\le\tau(A)\wedge\tau(B)\le\tau(A)\vee\tau(B)$
$\le\tau(A\vee B)\le\tau(A\oplus B)\le\tau(A)\oplus\tau(B).$

(ii) Furthermore, $$\bigotimes_{i=1}^{n}\tau(A_i)\le\tau(\bigotimes_{i=1}^{n}A_i)\le\tau(\bigwedge_{i=1}^{n}A_i)$$
$$\le\bigwedge_{i=1}^{n}\tau(A_i)\le\bigvee_{i=1}^{n}\tau(A_i)$$
$$\le\tau(\bigvee_{i=1}^{n}A_i)\le\tau(\bigoplus_{i=1}^{n}A_i)\le\bigoplus_{i=1}^{n}\tau(A_i).$$

Corollary 1. *Suppose that* $A_i\in F(S), i=1,2,\cdots,n$. *Then*

(i) $\tau(A_1\otimes A_2\otimes\cdots\otimes A_n)=1$ *if and only if* $\tau(A_i)=1, i=1,2,\cdots,n.$

(ii) $\tau(A_1\wedge A_2\wedge\cdots\wedge A_n)=1$ *if and only if* $\tau(A_i)=1, i=1,2,\cdots,n.$

Proposition 2. *Suppose that* $A_i\in F(S), i=1,2,\cdots,n$. *Then*

(i) $\tau(A_1\oplus A_2\oplus\cdots\oplus A_n)=0$ *if and only if* $\tau(A_i)=0, i=1,2,\cdots,n.$

(ii) $\tau(A_1\vee A_2\vee\cdots\vee A_n)=0$ *if and only if* $\tau(A_i)=0, i=1,2,\cdots,n.$

Corollary 2. *Suppose that* $A_i\in F(S), i=1,2,\cdots,n$. $k_1,k_2,\cdots,k_n\in\mathbf{N}$. *Then*

(i) $\tau(A_1^{k_1}\otimes A_2^{k_2}\otimes\cdots\otimes A_n^{k_n})=1$ *if and only if* $\tau(A_i)=1, i=1,2,\cdots,n.$

(ii) $\tau(k_1A_1\oplus k_2A_2\oplus\cdots\oplus k_nA_n)=0$ *if and only if* $\tau(A_i)=0, i=1,2,\cdots,n.$

where $A_i^{k_i}=\underbrace{A_i\otimes A_i\otimes\cdots\otimes A_i}_{k_i},\quad k_iA_i=\underbrace{A_i\oplus A_i\oplus\cdots\oplus A_i}_{k_i}.$

In **Łuk** weak deduction theorem[21] holds, i.e., if $\Gamma\cup\{A\}\vdash B$, then there exists $k\in\mathbf{N}$, such that $\Gamma\vdash A^k\to B$, where $A^k=\underbrace{A\otimes A\otimes\cdots\otimes A}_{k}$.

By weak deduction theorem and preceding results, it is easy to verify following theorem.

Theorem 3. *Let* $\Gamma\subset F(S)$ *be a theory. If* $\forall A\in\Gamma$, $\tau(A)=1$, *then* $\forall B\in D(\Gamma)$, $\tau(B)=1$.

Following results are obvious.

Proposition 3. *Suppose that* $A,B\in F(S)$.

(i) If A *is a refutable formula,i.e.,* $\vdash\neg A$, *then* $\tau(A)=0$.

(ii) $\tau(A)\oplus\tau(\neg A)=1, \tau(A)\otimes\tau(\neg A)=0.$

(iii) If $\tau(A\to B)=1$, *then* $\tau(A)\le\tau(B)$.

(iv) $\tau(A\to B)\le\tau(A)\to\tau(B)$.

Now gives out representations of $\wedge$ and $\vee$ under L-valued truth degree.

Theorem 4. *Suppose that* $A, B \in F(S)$. *Then*

$$\tau(A \wedge B) = \tau(A) \otimes \tau(A \to B). \tag{4}$$

Proof. Considering the fact that $A \oplus (A \to B) \sim \neg A \to (\neg B \to \neg A)$ is a theorem, we have $\tau(A \oplus (A \to B)) = 1$. Then by Theorem 1

$$\begin{aligned}\tau(A) \otimes \tau(A \to B) &= \tau(A \oplus (A \to B)) \otimes \tau(A \otimes (A \to B)) \\ &= \tau(A \otimes (A \to B)) \\ &= \tau(A \wedge B).\end{aligned}$$

Corollary 3. *Suppose that* $A, B \in F(S)$. *Then*

$$\tau(A \vee B) = \tau(A \to B) \to \tau(B). \tag{5}$$

Followings are form of L-valued truth degree under inference rules MP and HS.

Proposition 4. *Suppose that* $A, B \in F(S)$.
(i) If $\tau(A) \geq \alpha, \tau(A \to B) \geq \beta$, *then* $\tau(B) \geq \alpha \otimes \beta$.
(ii) If $\tau(A \to B) \geq \alpha, \tau(B \to C) \geq \beta$, *then* $\tau(A \to C) \geq \alpha \otimes \beta$.

Proof. (i) If $\tau(A) \geq \alpha, \tau(A \to B) \geq \beta$, then it is obtained that $\tau(B) \geq \tau(A \wedge B) = \tau(A) \otimes \tau(A \to B) \geq \alpha \otimes \beta$ from $\vdash (A \wedge B) \to B$ and equation (4).

(ii) is the direct result of (i).

Corollary 4. *Suppose that* $A, B \in F(S)$.
(i) If $\tau(A) = 1, \tau(A \to B) = 1$, *then* $\tau(B) = 1$.
(ii) If $\tau(A \to B) = 1, \tau(B \to C) = 1$, *then* $\tau(A \to C) = 1$.

Proposition 5. *Suppose that* $A, B \in F(S)$. *Then*
(i) $\tau(A \vee B) \otimes \tau(A \wedge B) = \tau(A) \otimes \tau(B)$.
(ii) $\tau(A \vee B) \oplus \tau(A \wedge B) = \tau(A) \oplus \tau(B)$.

Proof. (i) It is easy to verify that $B \sim (A \to B) \wedge B$ holds in **Łuk**. It is inferred from equations (4) and (5) that

$$\begin{aligned}\tau(A \vee B) \otimes \tau(A \wedge B) &= (\tau(A \to B) \to \tau(B)) \otimes (\tau(A) \otimes \tau(A \to B)) \\ &= \tau(A) \otimes [\tau(A \to B) \otimes (\tau(A \to B) \to \tau(B))] \\ &= \tau(A) \otimes \tau((A \to B) \wedge B) \\ &= \tau(A) \otimes \tau(B).\end{aligned}$$

(ii) is the direct result of (i).

Proposition 6. *Suppose that* $A, B \in F(S)$. *Then*

$$\tau(A) \otimes \tau(A \to B) = \tau(B) \otimes \tau(B \to A). \tag{6}$$

Proposition 7. *Suppose that* $A, B \in F(S)$. *Then*

$$\tau(A \to B) \to \tau(B) = \tau(B \to A) \to \tau(A). \tag{7}$$

Finally, two examples of L-valued truth degree are given as follows.

Example 1. Let $L = [F]$ be a Lindenbaum algebra of **Łuk**. Then L is a MV-algebra [2]. $\forall A \in F(S)$, put $\tau(A) = [A]$. Then it is easy to verify that τ is a truth degree on L in **Łuk**.

Example 2. Let L be a MV-algebra. Put $\upsilon : F(S) \to L$, and $\upsilon(\neg A) = (\upsilon(A))', \upsilon(A \to B) = \upsilon(A) \to \upsilon(B)$. Then it is easy to verify that υ is a truth degree on L in **Łuk**.

4 Conclusion

In this study, lattice-valued truth degree of formula is proposed in Łukasiewicz propositional fuzzy logic **Łuk**. Based on special properties of MV-algebra, it is easy to apply present method to lattice-valued proposition logic. Otherwise, lattice-valued consistent degree, lattice-valued resemblance degree between two formulae and so on can be studied based on lattice-valued truth degree. And other logic systems, such as propositional fuzzy logic $\mathcal{L}^*$, are also considered similarly.

Acknowledgements. The paper is supported by the Natural Science Foundation of Fujian Province of China (A0810014).

References

1. Wang, G.J., Fu, L., Song, J.S.: Theory of truth degrees of propositions in two-valued logic. Science in China (Series A) 45(9), 1106–1116 (2002)
2. Wang, G.J.: Introduction to mathematical logic and resolution principle, 2nd edn. Science Press, Beijing (2007) (in Chinese)
3. Li, B.J., Wang, G.J.: Theory of truth degrees of formulae in Łukasiwicz n-valued propositional logic and limit theorem. Science In China(Series F) 48(6), 727–736 (2005)
4. Li, J., Li, S.P., Xia, Y.F.: Theory of truth degrees of propositions in n-valued Łukasiwicz logic. Acta Mathematica Sinica 47(4), 769–780 (2004) (in Chinese)
5. Ru, Y.M., Wang, G.J.: Distribution of propositional truth degree in some 3-value propositional logic systems. Journal of Shaanxi Normal University (Natural Science Edition) 33(4), 6–10 (2005) (in Chinese)
6. Wang, G.J., Qin, X.Y., Zhou, X.N.: Theory of quasi-truth degrees of formulae in two-valued predicate logic. Journal of Shaanxi Normal University (Natural Science Edition) 33(1), 1–6 (2005) (in Chinese)
7. Wang, G.J., Leung, Y.: Integrated semantic and logic metric spaces. Fuzzy Sets and Systems 136, 71–91 (2003)
8. Wang, G.J., Duan, Q.L.: Theory of (n) truth degrees of formulae in modal logic and a consistency theorem. Science In China (Series F) 52(1), 70–83 (2009)
9. Wang, G.J., Zhang, W.X.: Consistency degrees of finite theories in Łukasiwicz propositional fuzzy logic. Fuzzy Sets and Systems 149, 275–284 (2005)

10. Zhou, X.N., Wang, G.J.: Consistency degrees of theories in some systems of propositional fuzzy logic. Fuzzy Sets and Systems 152, 321–331 (2005)
11. Zhou, H.J., Wang, G.J.: A new theory consistency index based on deduction theorems in several logic system. Fuzzy Sets and Systems 157, 427–443 (2006)
12. Wu, H.B., Wen, Q.M.: Calculus similar degree and pseudometric in calculus semantics. Journal of Shaanxi Normal University (Natural Science Edition) 28(3), 15–19 (2000) (in Chinese)
13. Wang, G.J.: Quantified logic (I). Journal of Engineering Mathematics 23(2), 191–215 (2006) (in Chinese)
14. Zhang, D.X., Li, L.F.: Syntactic graded method of two-valued propositional logic formulae. Acta Electronica Sinica 36(2), 325–330 (2008) (in Chinese)
15. Zhang, D.X., Chen, S.L.: A standard of formula truth degree in Łukasiewicz propositional fuzzy logic system. The Proceedings of the First International Conference on Quantitative Logic and Quantification of Software 1, 221–228 (2009)
16. Chang, C.C.: Algebraic analysis of many valued logics. Trans. Amer. Math. Soc. (88), 467–490 (1958)
17. Xu, Y., Qin, K.Y., Liu, J., Song, Z.M.: L-valued propositional logic L_{vpl}. Information Sciences 114, 205–235 (1999)
18. Xu, Y., Ruan, D., Qin, K.Y., Liu, J.: Lattice-valued logic. Springer, Berlin (2003)
19. Xu, Y.: Lattice implication algebra. Journal of Southwest Jiaotong University 28(1), 20–26 (1993)
20. Wang, G.J.: MV-algebra,BL-algebra,R_0-algebra and many-valued logic. Fuzzy Systems and Mathematics 16(2), 1–15 (2002)
21. Hájek, P.: Metamathematics of fuzzy logic. Kluwer Academic Publishers, Dordrecht (1998)

Decomposition Theorems and Representation Theorems on the IVIFS

Xue-Hai Yuan and Hong-Xing Li

Faculty of Electronic Information and Electrical Engineering,
Dalian University of Technology, Dalian 116024, China
e-mail: yuanxuehai@yahoo.com.cn, lihx@dlut.edu.cn

Abstract. In this paper, we established decomposition theorems and representation theorems of interval-valued intuitionistic fuzzy sets(IVIFS) by use of cut sets of interval-valued intuitionistic fuzzy sets. We have shown that each kind of cut sets corresponds to two kinds of decomposition theorems and representation theorems, thus eight kinds of decomposition theorems and representation theorems on interval-valued intuitionistic fuzzy sets are obtained, respectively. These results provide a fundamental theory for the research of interval-valued intuitionistic fuzzy sets.

Keywords: Interval-valued intuitionistic fuzzy set (IVIFS), cut set, decomposition theorem, representation theorem.

1 Introduction

Since the concept of fuzzy sets is introduced by Zadeh in 1965 [1], the theories of fuzzy sets and fuzzy systems are developed rapidly. As is well known, the cut set of fuzzy sets is an important concept in theory of fuzzy sets and systems, which plays an significant role in fuzzy topology [2,3], fuzzy algebra [4,5], fuzzy measure and fuzzy analysis [6-10], fuzzy optimization and decision making [11,12], fuzzy reasoning [13,14], fuzzy logic [15], and so on. The cut sets are the bridge connecting the fuzzy sets and classical sets. Based on it, the decomposition theorems and representation theorems can be established [16]. The cut sets on fuzzy sets are described in [17] by using the neighborhood relations between fuzzy point and fuzzy set. It is pointed that there are four kinds of definition of cut sets on fuzzy sets, each of which has similar properties. Also, the decomposition theorems and representation theorems can be established based on each kind of cut set. In [18],the interval-valued level cut sets of Zadeh fuzzy sets are presented and the decomposition theorems and representation theorems based on the interval-valued level cut sets are established.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 339–349.
springerlink.com

With the development of the theory on fuzzy sets, Goguen introduced L-fuzzy sets as an extension of Zadeh fuzzy sets in 1967[19]. Since then, many L-fuzzy sets are put forward, such as interval-valued fuzzy sets [20], intuitionistic fuzzy sets [21], interval-valued intuitionistic fuzzy sets [22],three-dimensional fuzzy sets[23] ,n-dimensional fuzzy sets[24] and type-2 fuzzy sets [25], and so on. Four new kind of cut sets of intuitionistic fuzzy sets are defined as 3-valued fuzzy sets and the decomposition theorems and representation theorems are established in [26]. In [27], the cut sets on interval-valued intutionistic fuzzy sets are defined as 5-valued fuzzy sets and we have shown that the cut sets of interval-valued intuitionistic fuzzy sets have similar properties with the cut sets of fuzzy sets and intuitionistic fuzzy sets. In [28], the theory of intuitionistic fuzzy subgroup is estabilished by use of the cut sets of intuitionistic fuzzy sets presented in [26]. However , the decomposition theorems and representation theorems of interval-valued intuitionistic fuzzy sets based on these cut sets have not been obtained so far.

In this paper, by use of the concept of cut sets of interval-valued intuitionistic fuzzy sets, we established decomposition theorems and representation theorems of interval-valued intuitionistic fuzzy sets. The rest of this paper is organized as follows: we first provide the preliminaries in section 2. In section 3 and section 4, we establish eight kinds of decomposition theorems and representation theorems on interval-valued intuitionistic fuzzy sets respectively.

2 Preliminary

Definition 1. [22] *Let X be a set and $M_A \subset [0,1], N_A \subset [0.1]$ be two closed intervals. If $supM_A(x)+supN_A(x) \leq 1, \forall x \in X$, then $A = (X, M_A, N_A)$ is called an interval-valued intuitionistic fuzzy sets over X. We let $\mathcal{IVIF}(X)$ denote the set of interval-valued intuitionistic fuzzy sets over X.*

Let $M_A(x) = [\mu_A^-(x), \mu_A^+(x)], N_A(x) = [\nu_A^-(x), \nu_A^+(x)]$. Then A can be denoted as $A = (X, [\mu_A^-(x), \mu_A^+(x)], [\nu_A^-(x), \nu_A^+(x)])$, where $0 \leq \mu_A^+(x)+\nu_A^+(x) \leq 1, \forall x \in X$.

For $A, B, A_t(t \in T) \in \mathcal{IVIF}(X), x \in X$, we set:

(1) $A \subset B \Leftrightarrow \mu_A^-(x) \leq \mu_B^-(x), \mu_A^+(x) \leq \mu_B^+(x)$ and $\nu_A^-(x) \geq \nu_B^-(x), \nu_A^+(x) \geq \nu_B^+(x), \forall x \in X$.

(2) $\underset{t\in T}{\cup} A_t = C = (X, [\mu_C^-(x), \mu_C^+(x)], [\nu_C^-(x), \nu_C^+(x)])$, where $\mu_C^-(x) = \underset{t\in T}{\vee} \mu_{A_t}^-(x)$, $\mu_C^+(x) = \underset{t\in T}{\vee} \mu_{A_t}^+(x)$, $\nu_C^-(x) = \underset{t\in T}{\wedge} \nu_{A_t}^-(x), \nu_C^+(x) = \underset{t\in T}{\wedge} \nu_{A_t}^+(x)$.

(3) $\underset{t\in T}{\cap} A_t = D = (X, [\mu_D^-(x), \mu_D^+(x)], [\nu_D^-(x), \nu_D^+(x)])$, where $\mu_D^-(x) = \underset{t\in T}{\wedge} \mu_{A_t}^-(x)$, $\mu_D^+(x) = \underset{t\in T}{\wedge} \mu_{A_t}^+(x)$, $\nu_D^-(x) = \underset{t\in T}{\vee} \nu_{A_t}^-(x), \nu_D^+(x) = \underset{t\in T}{\vee} \nu_{A_t}^+(x)$.

(4) $A^c = (X, [\nu_A^-(x), \nu_A^+(x)], [\mu_A^-(x), \mu_A^+(x)])$.

The definitions of cut sets on interval-valued intuitionistic fuzzy sets are given as follows:

Let $A = (X, [\mu_A^-(x), \mu_A^+(x)], [\nu_A^-(x), \nu_A^+(x)]) \in \mathcal{IVIF}(X), \lambda \in [0,1]$.

Definition 2. [27] *(1) If*

$$A_\lambda(x) = \begin{cases} 1, \mu_A^-(x) \geq \lambda, \\ \frac{3}{4},\ \mu_A^-(x) < \lambda \leq \mu_A^+(x), \\ \frac{1}{2},\ \mu_A^+(x) < \lambda \leq 1 - \nu_A^+(x), \\ \frac{1}{4},\ 1 - \nu_A^+(x) < \lambda \leq 1 - \nu_A^-(x), \\ 0,\ \lambda > 1 - \nu_A^-(x), \end{cases}$$

$$A_{\underline{\lambda}}(x) = \begin{cases} 1,\ \mu_A^-(x) > \lambda, \\ \frac{3}{4},\ \mu_A^-(x) \leq \lambda < \mu_A^+(x), \\ \frac{1}{2},\ \mu_A^+(x) \leq \lambda < 1 - \nu_A^+(x), \\ \frac{1}{4},\ 1 - \nu_A^+(x) \leq \lambda < 1 - \nu_A^-(x), \\ 0,\ \lambda \geq 1 - \nu_A^-(x), \end{cases}$$

then A_λ and $A_{\underline{\lambda}}$ are called $\lambda-$upper cut set and $\lambda-$strong upper cut set of A respectively.

(2) If

$$A^\lambda(x) = \begin{cases} 1,\ \nu_A^-(x) \geq \lambda, \\ \frac{3}{4},\ \nu_A^-(x) < \lambda \leq \nu_A^+(x), \\ \frac{1}{2},\ \nu_A^+(x) < \lambda \leq 1 - \mu_A^+(x), \\ \frac{1}{4},\ 1 - \mu_A^+(x) < \lambda \leq 1 - \mu_A^-(x), \\ 0,\ \lambda > 1 - \mu_A^-(x), \end{cases}$$

$$A^{\underline{\lambda}}(x) = \begin{cases} 1, \nu_A^-(x) > \lambda, \\ \frac{3}{4}, \nu_A^-(x) \leq \lambda < \nu_A^+(x), \\ \frac{1}{2}, \nu_A^+(x) \leq \lambda < 1 - \mu_A^+(x), \\ \frac{1}{4}, 1 - \mu_A^+(x) \leq \lambda < 1 - \mu_A^-(x), \\ 0, \lambda \geq 1 - \mu_A^-(x), \end{cases}$$

then A^λ and $A^{\underline{\lambda}}$ are called $\lambda-$lower cut set and $\lambda-$strong lower cut set of A respectively.

(3) If

$$A_{[\lambda]}(x) = \begin{cases} 1,\ \lambda + \mu_A^-(x) \geq 1, \\ \frac{3}{4},\ \mu_A^-(x) < 1 - \lambda \leq \mu_A^+(x), \\ \frac{1}{2},\ \mu_A^+(x) < 1 - \lambda \leq 1 - \nu_A^+(x), \\ \frac{1}{4},\ \nu_A^-(x) \leq \lambda < \nu_A^+(x), \\ 0,\ \nu_A^-(x) > \lambda, \end{cases}$$

$$A_{[\underline{\lambda}]}(x) = \begin{cases} 1, \lambda + \mu_A^-(x) > 1, \\ \frac{3}{4}, \mu_A^-(x) \leq 1 - \lambda < \mu_A^+(x), \\ \frac{1}{2}, \mu_A^+(x) \leq 1 - \lambda < 1 - \nu_A^+(x), \\ \frac{1}{4}, \nu_A^-(x) < \lambda \leq \nu_A^+(x), \\ 0, \nu_A^-(x) \geq \lambda. \end{cases}$$

then $A_{[\lambda]}$ and $A_{[\underline{\lambda}]}$ are called $\lambda-$upper Q-cut set and $\lambda-$strong upper Q-cut set of A respectively.

(4) If

$$A^{[\lambda]}(x)=\begin{cases}1, \lambda+\nu_A^-(x)\geq 1,\\ \frac{3}{4},\ \nu_A^-(x)<1-\lambda\leq \nu_A^+(x),\\ \frac{1}{2},\ \nu_A^+(x)<1-\lambda\leq 1-\mu_A^+(x),\\ \frac{1}{4},\ \mu_A^-(x)\leq\lambda<\mu_A^+(x),\\ 0,\ \mu_A^-(x)>\lambda,\end{cases}$$

$$A^{[\underline{\lambda}]}(x)=\begin{cases}1,\ \lambda+\nu_A^-(x)>1,\\ \frac{3}{4},\ \nu_A^-(x)\leq 1-\lambda< \nu_A^+(x),\\ \frac{1}{2},\ \nu_A^+(x)\leq 1-\lambda< 1-\mu_A^+(x),\\ \frac{1}{4},\ \mu_A^-(x)<\lambda\leq\mu_A^+(x),\\ 0,\mu_A^-(x)\geq\lambda,\end{cases}$$

then $A^{[\lambda]}$ and $A^{[\underline{\lambda}]}$ are called $\lambda-$lower Q-cut set and $\lambda-$strong lower Q- cut set of A respectively.

Let $A, B, A_t(t\in T)\in\mathcal{IVIF}(X), \lambda, \lambda_t(t\in T)\in[0,1], a=\bigwedge\limits_{t\in T}\lambda_t, b=\bigvee\limits_{t\in T}\lambda_t$, then we have the following properties:

Property 1. [27] (1) $A_{\underline{\lambda}}\subset A_\lambda$; $A\subset B\Rightarrow A_\lambda\subset B_\lambda, A_{\underline{\lambda}}\subset B_{\underline{\lambda}}$;
(2) $\lambda_1<\lambda_2\Rightarrow A_{\lambda_1}\supset A_{\lambda_2}, A_{\underline{\lambda_1}}\supset A_{\underline{\lambda_2}}, A_{\underline{\lambda_1}}\supset A_{\lambda_2}$.
(3) $(A\cup B)_\lambda=A_\lambda\cup B_\lambda, (A\cup B)_{\underline{\lambda}}=A_{\underline{\lambda}}\cup B_{\underline{\lambda}}$,
$(A\cap B)_\lambda=A_\lambda\cap B_\lambda, (A\cap B)_{\underline{\lambda}}=A_{\underline{\lambda}}\cap B_{\underline{\lambda}}$.
(4) $(A^c)_\lambda=(A_{\underline{1-\lambda}})^c, (A^c)_{\underline{\lambda}}=(A_{1-\lambda})^c$.
(5) $(\bigcup\limits_{t\in T}A_t)_\lambda\supset\bigcup\limits_{t\in T}(A_t)_\lambda, (\bigcup\limits_{t\in T}A_t)_{\underline{\lambda}}=\bigcup\limits_{t\in T}(A_t)_{\underline{\lambda}}$,
$(\bigcap\limits_{t\in T}A_t)_\lambda=\bigcap\limits_{t\in T}(A_t)_\lambda, (\bigcap\limits_{t\in T}A_t)_{\underline{\lambda}}\subset\bigcap\limits_{t\in T}(A_t)_{\underline{\lambda}}$;
(6) $\bigcup\limits_{t\in T}A_{\lambda_t}\subset A_a, \bigcap\limits_{t\in T}A_{\lambda_t}=A_b, \bigcup\limits_{t\in T}A_{\underline{\lambda_t}}=A_{\underline{a}}, \bigcap\limits_{t\in T}A_{\underline{\lambda_t}}\supset A_{\underline{b}}$;
(7) $A_{\underline{1}}=\emptyset, A_0=X$.

Property 2. [27] (1) $A^{\underline{\lambda}}\subset A^\lambda$;$A\subset B\Rightarrow B^\lambda\subset A^\lambda, B^{\underline{\lambda}}\subset A^{\underline{\lambda}}$;
(2) $\lambda_1<\lambda_2\Rightarrow A^{\lambda_1}\supset A^{\lambda_2}, A^{\underline{\lambda_1}}\supset A^{\underline{\lambda_2}}, A^{\underline{\lambda_2}}\supset A^{\lambda_1}$;
(3) $(A\cup B)^\lambda=A^\lambda\cap B^\lambda, (A\cup B)^{\underline{\lambda}}=A^{\underline{\lambda}}\cap B^{\underline{\lambda}}$,
$(A\cap B)^\lambda=A^\lambda\cup B^\lambda, (A\cap B)^{\underline{\lambda}}=A^{\underline{\lambda}}\cup B^{\underline{\lambda}}$;
(4) $(A^c)^\lambda=(A^{\underline{1-\lambda}})^c, (A^c)^{\underline{\lambda}}=(A^{1-\lambda})^c$;
(5) $(\bigcup\limits_{t\in T}A_t)^\lambda=\bigcap\limits_{t\in T}(A_t)^\lambda, (\bigcup\limits_{t\in T}A_t)^{\underline{\lambda}}\subset\bigcap\limits_{t\in T}(A_t)^{\underline{\lambda}}$,
$(\bigcap\limits_{t\in T}A_t)^\lambda\supset\bigcup\limits_{t\in T}(A_t)^\lambda, (\bigcap\limits_{t\in T}A_t)^{\underline{\lambda}}=\bigcup\limits_{t\in T}(A_t)^{\underline{\lambda}}$;
(6) $\bigcup\limits_{t\in T}A^{\lambda_t}\subset A^a, \bigcap\limits_{t\in T}A^{\lambda_t}=A^b, \bigcup\limits_{t\in T}A^{\underline{\lambda_t}}=A^{\underline{a}}, \bigcap\limits_{t\in T}A^{\underline{\lambda_t}}\supset A^{\underline{b}}$;
(7) $A^{\underline{0}}=\emptyset, A^1=X$.

Property 3. [27] (1) $A_{[\underline{\lambda}]}\subset A_{[\lambda]}$;$A\subset B\Rightarrow A_{[\lambda]}\subset B_{[\lambda]}, A_{[\underline{\lambda}]}\subset B_{[\underline{\lambda}]}$;
(2) $\lambda_1<\lambda_2\Rightarrow A_{[\lambda_1]}\subset A_{[\lambda_2]}, A_{[\underline{\lambda_1}]}\subset A_{[\underline{\lambda_2}]}, A_{[\lambda_1]}\subset A_{[\underline{\lambda_2}]}$;
(3) $(A\cup B)_{[\lambda]}=A_{[\lambda]}\cup B_{[\lambda]}, (A\cup B)_{[\underline{\lambda}]}=A_{[\underline{\lambda}]}\cup B_{[\underline{\lambda}]}$,
$(A\cap B)_{[\lambda]}=A_{[\lambda]}\cap B_{[\lambda]}, (A\cap B)_{[\underline{\lambda}]}=A_{[\underline{\lambda}]}\cap B_{[\underline{\lambda}]}$;
(4) $(A^c)_{[\lambda]}=(A_{[\underline{1-\lambda}]})^c, (A^c)_{[\underline{\lambda}]}=(A_{[1-\lambda]})^c$;

(5) $(\bigcup_{t\in T} A_t)_{[\lambda]} \supset \bigcup_{t\in T}(A_t)_{[\lambda]}, (\bigcup_{t\in T} A_t)_{[\underline{\lambda}]} = \bigcup_{t\in T}(A_t)_{[\underline{\lambda}]}$,
$(\bigcap_{t\in T} A_t)_{[\lambda]} = \bigcap_{t\in T}(A_t)_{[\lambda]}, (\bigcap_{t\in T} A_t)_{[\underline{\lambda}]} \subset \bigcap_{t\in T}(A_t)_{[\underline{\lambda}]}$;
(6) $\bigcup_{t\in T} A_{[\lambda_t]} \subset A_{[b]}, \bigcap_{t\in T} A_{[\lambda_t]} = A_{[a]}, \bigcup_{t\in T} A_{[\underline{\lambda_t}]} = A_{[\underline{b}]}, \bigcap_{t\in T} A_{[\underline{\lambda_t}]} \supset A_{[\underline{a}]}$;
(7) $A_{[\underline{0}]} = \emptyset, A_{[1]} = X$.

Property 4. [27] (1) $A^{[\underline{\lambda}]} \subset A^{[\lambda]}$; $A \subset B \Rightarrow B^{[\lambda]} \subset A^{[\lambda]}, B^{[\underline{\lambda}]} \subset A^{[\underline{\lambda}]}$;
(2) $\lambda_1 < \lambda_2 \Rightarrow A^{[\lambda_1]} \subset A^{[\lambda_2]}, A^{[\underline{\lambda_1}]} \subset A^{[\underline{\lambda_2}]}$, $A^{[\lambda_1]} \subset A^{[\underline{\lambda_2}]}$;
(3) $(A\cup B)^{[\lambda]} = A^{[\lambda]} \cap B^{[\lambda]}, (A\cup B)^{[\underline{\lambda}]} = A^{[\underline{\lambda}]} \cap B^{[\underline{\lambda}]}$,
$(A\cap B)^{[\lambda]} = A^{[\lambda]} \cup B^{[\lambda]}, (A\cap B)^{[\underline{\lambda}]} = A^{[\underline{\lambda}]} \cup B^{[\underline{\lambda}]}$;
(4) $(A^c)^{[\lambda]} = (A^{[\underline{1-\lambda}]})^c, (A^c)^{[\underline{\lambda}]} = (A^{[1-\lambda]})^c$;
(5) $(\bigcup_{t\in T} A_t)^{[\lambda]} = \bigcap_{t\in T}(A_t)^{[\lambda]}, (\bigcup_{t\in T} A_t)^{[\underline{\lambda}]} \subset \bigcap_{t\in T}(A_t)^{[\underline{\lambda}]}$,
$(\bigcap_{t\in T} A_t)^{[\lambda]} \supset \bigcup_{t\in T}(A_t)^{[\lambda]}, (\bigcap_{t\in T} A_t)^{[\underline{\lambda}]} = \bigcup_{t\in T}(A_t)^{[\underline{\lambda}]}$;
(6) $\bigcup_{t\in T} A^{[\lambda_t]} \subset A^{[b]}, \bigcap_{t\in T} A^{[\lambda_t]} = A^{[a]}, \bigcup_{t\in T} A^{[\underline{\lambda_t}]} = A^{[\underline{b}]}, \bigcap_{t\in T} A^{[\underline{\lambda_t}]} \supset A^{[\underline{a}]}$;
(7) $A^{[\underline{1}]} = \emptyset, A^{[0]} = X$.

3 Decomposition Theorems

Let $5^X = \{f | f : X \to \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$ is a mapping$\}$, We give the following definition:

Definition 3. *Let mappings* $f_i : [0,1] \times 5^X \to \mathcal{IVIF}(X)$ $(\lambda, A) \mapsto f_i(\lambda, A)(i = 1, 2, \cdots, 8)$ *and*

$$f_1(\lambda, A)(x) == \begin{cases} ([0,0],[1,1]), & A(x) = 0, \\ ([0,0],[1-\lambda,1]), & A(x) = \frac{1}{4}, \\ ([0,0],[1-\lambda,1-\lambda]), & A(x) = \frac{1}{2}, \\ ([0,\lambda],[1-\lambda,1-\lambda]), & A(x) = \frac{3}{4}, \\ ([\lambda,\lambda],[1-\lambda,1-\lambda]), & A(x) = 1, \end{cases}$$

$$f_2(\lambda, A)(x) == \begin{cases} ([\lambda,\lambda],[1-\lambda,1-\lambda]), A(x) = 0, \\ ([\lambda,\lambda],[0,1-\lambda]), A(x) = \frac{1}{4}, \\ ([\lambda,\lambda],[0,0]), A(x) = \frac{1}{2}, \\ ([\lambda,1],[0,0]), A(x) = \frac{3}{4}, \\ ([1,1],[0,0]), A(x) = 1, \end{cases}$$

$$f_3(\lambda, A)(x) == \begin{cases} ([1-\lambda,1-\lambda],[\lambda,\lambda]), A(x) = 0, \\ ([0,1-\lambda],[\lambda,\lambda]), A(x) = \frac{1}{4}, \\ ([0,0],[\lambda,\lambda]), A(x) = \frac{1}{2}, \\ ([0,0],[\lambda,1]), A(x) = \frac{3}{4}, \\ ([0,0],[1,1]), A(x) = 1, \end{cases}$$

$$f_4(\lambda,A)(x) == \begin{cases} ([1,1],[0,0]), A(x)=0, \\ ([1-\lambda,1],[0,0]), A(x)=\frac{1}{4}, \\ ([1-\lambda,1-\lambda],[0,0]), A(x)=\frac{1}{2}, \\ ([1-\lambda,1-\lambda],[0,\lambda]), A(x)=\frac{3}{4}, \\ ([1-\lambda,1-\lambda],[\lambda,\lambda]), A(x)=1, \end{cases}$$

$$f_5(\lambda,A)(x) == \begin{cases} ([0,0],[1,1]), A(x)=0, \\ ([0,0],[\lambda,1]), A(x)=\frac{1}{4}, \\ ([0,0],[\lambda,\lambda]), A(x)=\frac{1}{2}, \\ ([0,1-\lambda],[\lambda,\lambda]), A(x)=\frac{3}{4}, \\ ([1-\lambda,1-\lambda],[\lambda,\lambda]), A(x)=1, \end{cases}$$

$$f_6(\lambda,A)(x) == \begin{cases} ([1-\lambda,1-\lambda],[\lambda,\lambda]), A(x)=0, \\ ([1-\lambda,1-\lambda],[0,\lambda]), A(x)=\frac{1}{4}, \\ ([1-\lambda,1-\lambda],[0,0]), A(x)=\frac{1}{2}, \\ ([1-\lambda,1],[0,0]), A(x)=\frac{3}{4}, \\ ([1,1],[0,0]), A(x)=1, \end{cases}$$

$$f_7(\lambda,A)(x) == \begin{cases} ([\lambda,\lambda],[1-\lambda,1-\lambda]), A(x)=0, \\ ([0,\lambda],[1-\lambda,1-\lambda]), A(x)=\frac{1}{4}, \\ ([0,0],[1-\lambda,1-\lambda]), A(x)=\frac{1}{2}, , \\ ([0,0],[1-\lambda,1]), A(x)=\frac{3}{4}, \\ ([0,0],[1,1]), A(x)=1, \end{cases}$$

$$f_8(\lambda,A)(x) == \begin{cases} ([1,1],[0,0]), A(x)=0, \\ ([\lambda,1],[0,0]), A(x)=\frac{1}{4}, \\ ([\lambda,\lambda],[0,0]), A(x)=\frac{1}{2}, \\ ([\lambda,\lambda],[0,1-\lambda]), A(x)=\frac{3}{4}, \\ ([\lambda,\lambda],[1-\lambda,1-\lambda]), A(x)=1. \end{cases}$$

Then we have the following decomposition theorems.

Theorem 1. *Let $A \in \mathcal{IVIF}(X)$. Then*

(1) $A = \bigcup\limits_{\lambda\in[0,1]} f_1(\lambda, A_\lambda) = \bigcap\limits_{\lambda\in[0,1]} f_2(\lambda, A_\lambda)$;

(2) $A = \bigcup\limits_{\lambda\in[0,1]} f_1(\lambda, A_{\underline{\lambda}}) = \bigcap\limits_{\lambda\in[0,1]} f_2(\lambda, A_{\underline{\lambda}})$.

(3) If the mapping $H : [0,1] \to 5^X$ *satisfies* $A_{\underline{\lambda}} \subset H(\lambda) \subset A_\lambda$, *then*

(i) $A = \bigcup\limits_{\lambda\in[0,1]} f_1(\lambda, H(\lambda)) = \bigcap\limits_{\lambda\in[0,1]} f_2(\lambda, H(\lambda))$;

(ii) $\lambda_1 < \lambda_2 \Rightarrow H(\lambda_1) \supset H(\lambda_2)$; *(iii)* $A_\lambda = \bigcap\limits_{\alpha<\lambda} H(\alpha), A_{\underline{\lambda}} = \bigcup\limits_{\alpha>\lambda} H(\alpha)$.

Proof. (1) $(\bigcup\limits_{\lambda\in[0,1]} f_1(\lambda, A_\lambda))(x) = \bigvee\limits_{\lambda\in[0,1]} f_1(\lambda, A_\lambda)(x) = ([\vee\{\lambda|A_\lambda(x) = 1\}, \vee\{\lambda|A_\lambda(x) \geq \frac{3}{4}\}], [\wedge\{1-\lambda|A_\lambda(x) \geq \frac{1}{4}\}, \wedge\{1-\lambda|A_\lambda(x) \geq \frac{1}{2}\}])$.

Since $\vee\{\lambda|A_\lambda(x) = 1\} = \vee\{\lambda|\mu_A^-(x) \geq \lambda\} = \mu_A^-(x)$, $\vee\{\lambda|A_\lambda(x) \geq \frac{3}{4}\} = \vee\{\lambda|\mu_A^+(x) \geq \lambda\} = \mu_A^+(x)$, $\wedge\{1-\lambda|A_\lambda(x) \geq \frac{1}{4}\} = \wedge\{1-\lambda|1-\nu_A^-(x) \geq \lambda\} = \wedge\{1-\lambda|1-\lambda \geq \nu_A^-(x)\} = \nu_A^-(x)$, $\wedge\{1-\lambda|A_\lambda(x) \geq \frac{1}{2}\} = \wedge\{1-\lambda|1-\nu_A^+(x) \geq \lambda\} = \wedge\{1-\lambda|1-\lambda \geq \nu_A^+(x)\} = \nu_A^+(x)$, so $(\bigcup_{\lambda\in[0,1]} f_1(\lambda, A_\lambda))(x) = ([\mu_A^-(x), \mu_A^+(x)], [\nu_A^-(x), \nu_A^+(x)]) = A(x)$. Thus, $A = (\bigcup_{\lambda\in[0,1]} f_1(\lambda, A_\lambda))$.

Similarly, $(\bigcap_{\lambda\in[0,1]} f_2(\lambda, A_\lambda))(x) = \bigwedge_{\lambda\in[0,1]} f_2(\lambda, A_\lambda)(x) = ([\wedge\{\lambda|A_\lambda(x) \leq \frac{3}{4}\}, \wedge\{\lambda|A_\lambda(x) \leq \frac{1}{2}\}], [\vee\{1-\lambda|A_\lambda(x) = 0\}, \vee\{1-\lambda|A_\lambda(x) \leq \frac{1}{4}\}])$.

Since $\wedge\{\lambda|A_\lambda(x) \leq \frac{3}{4}\} = \wedge\{\lambda|\mu_A^-(x) < \lambda\} = \mu_A^-(x)$, $\wedge\{\lambda|A_\lambda(x) \leq \frac{1}{2}\} = \wedge\{\lambda|\mu_A^+(x) < \lambda\} = \mu_A^+(x)$, $\vee\{1-\lambda|A_\lambda(x) = 0\} = \vee\{1-\lambda\|\lambda > 1-\nu_A^-(x)\} = \vee\{1-\lambda|\nu_A^-(x) > 1-\lambda\} = \nu_A^-(x)$, $\vee\{1-\lambda|A_\lambda(x) \leq \frac{1}{4}\} = \vee\{1-\lambda|\lambda > 1-\nu_A^+(x)\} = \vee\{1-\lambda|\nu_A^+(x) > 1-\lambda\} = \nu_A^+(x)$, so $(\bigcap_{\lambda\in[0,1]} f_2(\lambda, A_\lambda))(x) = ([\mu_A^-(x), \mu_A^+(x)], [\nu_A^-(x), \nu_A^+(x)]) = A(x)$. Thus, $A = (\bigcap_{\lambda\in[0,1]} f_2(\lambda, A_\lambda))$.

Proof of (2) is similar.

(3) (i) By $A_{\underline{\lambda}} \subset H(\lambda) \subset A_\lambda$, we have that

$f_1(\lambda, A_{\underline{\lambda}}) \subset f_1(\lambda, H(\lambda)) \subset f_1(\lambda, A_\lambda)$, $f_2(\lambda, A_{\underline{\lambda}}) \subset f_2(\lambda, H(\lambda)) \subset f_2(\lambda, A_\lambda)$.

Thus $A = \bigcup_{\lambda\in[0,1]} f_1(\lambda, A_{\underline{\lambda}}) \subset \bigcup_{\lambda\in[0,1]} f_1(\lambda, H(\lambda)) \subset \bigcup_{\lambda\in[0,1]} f_1(\lambda, A_\lambda) = A$,

$$A = \bigcap_{\lambda\in[0,1]} f_2(\lambda, A_{\underline{\lambda}}) \subset \bigcap_{\lambda\in[0,1]} f_2(\lambda, H(\lambda)) \subset \bigcap_{\lambda\in[0,1]} f_2(\lambda, A_\lambda) = A.$$

Therefore, $A = \bigcup_{\lambda\in[0,1]} f_1(\lambda, H(\lambda)) = \bigcap_{\lambda\in[0,1]} f_2(\lambda, H(\lambda))$.

(ii) $\lambda_1 < \lambda_2 \Rightarrow H(\lambda_1) \supset A_{\underline{\lambda_1}} \supset A_{\lambda_2} \supset H(\lambda_2)$.

(iii) $\alpha < \lambda \Rightarrow H(\alpha) \supset A_\alpha \supset A_\lambda \Rightarrow \bigcap_{\alpha>\lambda} H(\alpha) \supset A_\lambda$.

Since $A_\lambda = \bigcap_{\alpha>\lambda} A_{\underline{\alpha}} \supset \bigcap_{\alpha>\lambda} H(\alpha) \supset A_\lambda \supset A_\lambda$, so $A_\lambda = \bigcap_{\alpha<\lambda} H(\alpha)$.

Similarly, $A_{\underline{\lambda}} = \bigcup_{\alpha>\lambda} H(\alpha)$.

Remark 1. (1) From Theorem 1, we know that the decomposition theorems of interval-valued intuitionistic fuzzy sets based on upper cut sets have been established.

(2) If $f_1(\lambda, A)$ and $f_2(\lambda, A)$ are denoted as λA and $\lambda \circ A$ respectively, then we have that $A = \bigcup_{\lambda\in[0,1]} \lambda A = \bigcap_{\lambda\in[0,1]} \lambda \circ A$, which are consistent with the normal decomposition theorems of Zadeh fuzzy sets and intuitionistic fuzzy sets[17, 26].

Similarly, we have the following theorems:

Theorem 2. *Let $A \in \mathcal{IVIF}(X)$. Then*

(1) $A = \bigcup_{\lambda\in[0,1]} f_3(\lambda, A^\lambda) = \bigcap_{\lambda\in[0,1]} f_4(\lambda, A^\lambda)$;

(2) $A = \bigcup_{\lambda\in[0,1]} f_3(\lambda, A^{\underline{\lambda}}) = \bigcap_{\lambda\in[0,1]} f_4(\lambda, A^{\underline{\lambda}})$;

(3) If the mapping $H : [0,1] \to 5^X$ *satisfies* $A^{\underline{\lambda}} \subset H(\lambda) \subset A^\lambda$,*then*

(i) $A = \bigcup_{\lambda\in[0,1]} f_3(\lambda, H(\lambda)) = \bigcap_{\lambda\in[0,1]} f_4(\lambda, H(\lambda))$*;*
(ii) $\lambda_1 < \lambda_2 \Rightarrow H(\lambda_1) \subset H(\lambda_2)$*; (iii)* $A^\lambda = \bigcap_{\alpha<\lambda} H(\alpha), A^{\underline{\lambda}} = \bigcup_{\alpha>\lambda} H(\alpha)$.

Theorem 3. *Let* $A \in \mathcal{IVIF}(X)$*. Then*
(1) $A = \bigcup_{\lambda\in[0,1]} f_5(\lambda, A_{[\lambda]}) = \bigcap_{\lambda\in[0,1]} f_6(\lambda, A_{[\lambda]})$*;*
(2)) $A = \bigcup_{\lambda\in[0,1]} f_5(\lambda, A_{[\underline{\lambda}]}) = \bigcap_{\lambda\in[0,1]} f_6(\lambda, A_{[\underline{\lambda}]})$*;*
(3) If the mapping $H : [0,1] \to 5^X$ *satisfies* $A_{[\underline{\lambda}]} \subset H(\lambda) \subset A_{[\lambda]}$*, then*
(i) $A = \bigcup_{\lambda\in[0,1]} f_5(\lambda, H(\lambda)) = \bigcap_{\lambda\in[0,1]} f_6(\lambda, H(\lambda))$*;*
(ii) $\lambda_1 < \lambda_2 \Rightarrow H(\lambda_1) \subset H(\lambda_2)$*; (iii)* $A_{[\lambda]} = \bigcap_{\alpha>\lambda} H(\alpha), A_{[\underline{\lambda}]} = \bigcup_{\alpha<\lambda} H(\alpha)$.

Theorem 4. *Let* $A \in \mathcal{IVIF}(X)$*. Then*
(1) $A = \bigcup_{\lambda\in[0,1]} f_7(\lambda, A^{[\lambda]}) = \bigcap_{\lambda\in[0,1]} f_8(\lambda, A^{[\lambda]})$*;*
(2) $A = \bigcup_{\lambda\in[0,1]} f_7(\lambda, A^{[\underline{\lambda}]}) = \bigcap_{\lambda\in[0,1]} f_8(\lambda, A^{[\underline{\lambda}]})$*;*
(3) If the mapping $H : [0,1] \to 5^X$ *satisfies* $A^{[\underline{\lambda}]} \subset H(\lambda) \subset A^{[\lambda]}$*, then*
(i) $A = \bigcup_{\lambda\in[0,1]} f_7(\lambda, H(\lambda)) = \bigcap_{\lambda\in[0,1]} f_8(\lambda, H(\lambda))$*;*
(ii) $\lambda_1 < \lambda_2 \Rightarrow H(\lambda_1) \supset H(\lambda_2)$*; (iii)* $A^{[\lambda]} = \bigcap_{\alpha>\lambda} H(\alpha), A^{[\underline{\lambda}]} = \bigcup_{\alpha<\lambda} H(\alpha)$.

4 Representation Theorems

Definition 4. *Let* $H : [0,1] \to 5^X$ *be a mapping,*
(1) If $(\lambda_1 < \lambda_2 \Rightarrow H(\lambda_1) \supset H(\lambda_2))$*, then* H *is called inverse order nested set over* X*; (2) If* $(\lambda_1 < \lambda_2 \Rightarrow H(\lambda_1) \subset H(\lambda_2))$*, then* H *is called a order nested set over* X*.*

For example, $H_1(\lambda) = A_\lambda$ and $H_2(\lambda) = A^{[\lambda]}$ are inverse order nested set over X; $H_3(\lambda) = A^\lambda$ and $H_4(\lambda) = A_{[\lambda]}$ are order nested set over X.

Let $\mathcal{U}(X)$ and $\mathcal{V}(X)$ be sets of inverse order nested set over X and order nested set over X, respectively. We set operations in $\mathcal{U}(X)$ as follows:

(a) $H_1 \subset H_2 \Leftrightarrow H_1(\lambda) \subset H_2(\lambda), \forall\lambda \in [0.1]$; (b) $(\bigcup_{t\in T} H_t)(\lambda) = \bigcup_{t\in T} H_t(\lambda)$;
(c) $(\bigcap_{t\in T} H_t)(\lambda) = \bigcap_{t\in T} H_t(\lambda)$; (d) $(H^c)(\lambda) = (H(1-\lambda))^c$,

then $(\mathcal{U}(X), \bigcup, \bigcap, C)$ is a De Morgan algebra. Similarly, we set operations in $\mathcal{V}(X)$ as follows:

(a) $H_1 \subset H_2 \Leftrightarrow H_1(\lambda) \supset H_2(\lambda), \forall\lambda \in [0.1]$; (b) $(\bigcup_{t\in T} H_t)(\lambda) = \bigcap_{t\in T} H_t(\lambda)$;
(c) $(\bigcap_{t\in T} H_t)(\lambda) = \bigcup_{t\in T} H_t(\lambda)$; (d) $(H^c)(\lambda) = (H(1-\lambda))^c$,

then $(\mathcal{V}(X), \bigcup, \bigcap, C)$ is also a De Morgan algebra.

Let the mappings $T_i : \mathcal{U}(X) \to \mathcal{IVIF}(X)(i = 1, 2, 7, 8)$ satisfy:

$T_i(H) = \bigcup_{\lambda\in[0,1]} f_i(\lambda, H(\lambda))$ $(i = 1, 7)$, $T_i(H) = \bigcap_{\lambda\in[0,1]} f_i(\lambda, H(\lambda))$ $(i = 2, 8)$.

Similarly, let the mappings $T_i : \mathcal{V}(X) \to \mathcal{IVIF}(X)(i = 3, 4, 5, 6)$ satisfy:
$T_i(H) = \bigcup_{\lambda\in[0,1]} f_i(\lambda, H(\lambda))$ $(i = 3, 5)$, $T_i(H) = \bigcap_{\lambda\in[0,1]} f_i(\lambda, H(\lambda))$ $(i = 4, 6)$.

Then we have the following representation theorems:

Theorem 5. *(1)* $T_1(H) = T_2(H)$*; (2)* $T_1(T_2)$ *is surjection ;*
(3) $T_1(\bigcup_{t\in T} H_t) = \bigcup_{t\in T} T_1(H_t)$, $T_1(\bigcap_{t\in T} H_t) = \bigcap_{t\in T} T_1(H_t)$, $T_1(H^c) = (T_1(H))^c$.

Proof. :(1) Let $A = T_1(H)$, then
$\mu_A^-(x) = \vee\{\lambda|H(\lambda)(x) = 1\}$, $\mu_A^+(x) = \vee\{\lambda|H(\lambda)(x) \geq \frac{3}{4}\}$,
$\nu_A^-(x) = \wedge\{1-\lambda|H(\lambda)(x) \geq \frac{1}{4}\}$, $\nu_A^+(x) = \wedge\{1-\lambda|H(\lambda)(x) \geq \frac{1}{2}\}$.
We first show that $T_1(H)_{\underline{\lambda}} \subset H(\lambda) \subset T_1(H)_\lambda$.
In fact, when $H(\lambda)(x) = 1$, we have that $\mu_A^-(x) \geq \lambda$, then $T_1(H)_\lambda = 1$.
When $H(\lambda)(x) = \frac{3}{4}$, we have that $\mu_A^+(x) \geq \lambda$, then $T_1(H)_\lambda \geq \frac{3}{4}$.
When $H(\lambda)(x) = \frac{1}{2}$, we have that $\nu_A^+(x) \leq 1-\lambda$, then $T_1(H)_\lambda \geq \frac{1}{2}$.
When $H(\lambda)(x) = \frac{1}{4}$, we have that $\nu_A^-(x) \leq 1-\lambda$, then $T_1(H)_\lambda \geq \frac{1}{4}$.
Since $H(\lambda), T_1(H)_\lambda \in 5^X$, so $H(\lambda) \subset T_1(H)_\lambda$.
On the other hand, $T_1(H)_{\underline{\lambda}} = 1 \Rightarrow \mu_A^-(x) > \lambda \Rightarrow \exists\alpha > \lambda, H(\alpha)(x) = 1 \Rightarrow H(\lambda)(x) \geq H(\alpha)(x) = 1 \Rightarrow H(\lambda)(x) = 1$; $T_1(H)_{\underline{\lambda}} = \frac{3}{4} \Rightarrow \mu_A^+(x) > \lambda \Rightarrow \exists\alpha > \lambda, H(\alpha)(x) \geq \frac{3}{4} \Rightarrow H(\lambda)(x) \geq H(\alpha)(x) \geq \frac{3}{4}$; $T_1(H)_{\underline{\lambda}} = \frac{1}{2} \Rightarrow 1-\nu_A^+(x) > \lambda \Rightarrow \nu_A^+(x) < 1-\lambda \Rightarrow \exists\alpha, 1-\alpha < 1-\lambda, H(\alpha)(x) \geq \frac{1}{2} \Rightarrow \exists\alpha > \lambda, H(\lambda)(x) \geq H(\alpha)(x) \geq \frac{1}{2}$; $T_1(H)_{\underline{\lambda}} = \frac{1}{4} \Rightarrow 1-\nu_A^-(x) > \lambda \Rightarrow \nu_A^-(x) < 1-\lambda \Rightarrow \exists\alpha, 1-\alpha < 1-\lambda, H(\alpha)(x) \geq \frac{1}{4} \Rightarrow \exists\alpha > \lambda, H(\lambda)(x) \geq H(\alpha)(x) \geq \frac{1}{4}$. Since $H(\lambda), T_1(H)_{\underline{\lambda}} \in 5^X$, so $H(\lambda) \supset T_1(H)_\lambda$. Therefore, $T_1(H)_{\underline{\lambda}} \subset H(\lambda) \subset T_1(H)_\lambda$. By Theorem 1, we have that $T_1(H) = \bigcup_{\lambda\in[0,1]} f_1(\lambda, H(\lambda)) = \bigcap_{\lambda\in[0,1]} f_2(\lambda, H(\lambda)) = T_2(H)$.

(2) Let $A \in \mathcal{IVIF}(X)$ and $H(\lambda) = A_\lambda$. Then $T_1(H) = A$.

(3) By Theorem 1, we have that $T_1(H)_\lambda = \bigcap_{\alpha<\lambda} H(\alpha), T_1(H)_{\underline{\lambda}} = \bigcup_{\alpha>\lambda} H(\alpha)$.
Thus, $T_1(\bigcup_{t\in T} H_t)_{\underline{\lambda}} = \bigcup_{\alpha>\lambda}(\bigcup_{t\in T} H_t)(\alpha) = \bigcup_{\alpha>\lambda}\bigcup_{t\in T} H_t(\alpha) = \bigcup_{t\in T}\bigcup_{\alpha>\lambda} H_t(\alpha) = (\bigcup_{t\in T}(T_1(H_t))_{\underline{\lambda}} = (\bigcup_{t\in T} T_1(H_t))_{\underline{\lambda}}$. By Theorem 1, we have that $T_1(\bigcup_{t\in T} H_t) = \bigcup_{t\in T} T_1(H_t)$.
Similarly, $T_1(\bigcap_{t\in T} H_t)_\lambda = \bigcap_{\alpha<\lambda}(\bigcap_{t\in T} H_t)(\alpha) = \bigcap_{\alpha<\lambda}\bigcap_{t\in T} H_t(\alpha) = \bigcap_{t\in T}\bigcap_{\alpha<\lambda} H_t(\alpha) = \bigcap_{t\in T}(T_1(H_t))_\lambda = (\bigcap_{t\in T} T_1(H_t))_\lambda$; $T_1(H^c)_\lambda = \bigcap_{\alpha<\lambda} H^c(\alpha) = \bigcap_{\alpha<\lambda}(H(1-\alpha))^c = (\bigcup_{1-\alpha>1-\lambda} H(1-\alpha))^c = (\bigcup_{\alpha>1-\lambda} H(\alpha))^c = (T_1(H)_{\underline{1-\lambda}})^c = ((T_1(H))^c)_\lambda$. By Theorem 1, we have that $T_1(\bigcap_{t\in T} H_t) = \bigcap_{t\in T} T_1(H_t)$, $T_1(H^c) = (T_1(H))^c$.

Similarly, we have

Theorem 6. *(1)* $T_3(H) = T_4(H)$*; (2)* $T_3(T_4)$ *is surjection;*
(3) $T_3(\bigcup_{t\in T} H_t) = \bigcup_{t\in T} T_3(H_t), T_3(\bigcap_{t\in T} H_t) = \bigcap_{t\in T} T_3(H_t), T_3(H^c) = (T_3(H))^c$.

Theorem 7. *(1)* $T_5(H) = T_6(H)$*; (2)* $T_5(T_6)$ *is surjection;*
(3) $T_5(\bigcup_{t\in T} H_t) = \bigcap_{t\in T} T_5(H_t), T_5(\bigcap_{t\in T} H_t) = \bigcup_{t\in T} T_5(H_t), T_5(H^c) = (T_5(H))^c$.

Theorem 8. *(1)* $T_7(H) = T_8(H)$*; (2)* $T_7(T_8)$ *is surjection;*
(3) $T_7(\bigcup_{t\in T} H_t) = \bigcap_{t\in T} T_7(H_t), T_7(\bigcap_{t\in T} H_t) = \bigcup_{t\in T} T_7(H_t), T_7(H^c) = (T_7(H))^c$.

Remark 2. From Theorem 5-Theorem 8 we have known that the representation theorems of interval-valued intuitionistic fuzzy sets have been established.

Acknowledgements. This research was supported in part by National Natural Science Foundation of China (No.60774049) and National Natural Science Foundation of China (Major Research Plan, No.90818025).

References

1. Zadeh, L.A.: Fuzzy Sets. Information and Control 8, 352–378 (1965)
2. Wang, G.J.: L-fuzzy Topology Space Theory. Shanxi Normal University Press, Xian (1988) (in Chinese)
3. Liu, Y.M., Luo, M.K.: Fuzzy Topology. World Scientific Publishing, Singapore (1990)
4. Mordeson, J.N., Malik, D.S.: Fuzzy Commutative Algebra. World Scientific Publishing, Singapore (1998)
5. Mordeson, J.N., Bhutani, K.R., Rosenfeld, A.: Fuzzy Group Theory. Springer, New York (2005)
6. Zhang, G.Q.: Fuzzy Measure Theory. Guizhou Science and Technology Press, Guiyang (1994) (in Chinese)
7. Wu, C.X., Ma, M.: The Basis of Fuzzy Analysis. National Defence Industry Press, Beijing (1991) (in Chinese)
8. Bertoluzza, C., Solci, M., Capodieci, M.L.: Measure of a fuzzy set: the α-cut approach in the finite case. Fuzzy Sets and Systems 123(1), 93–102 (2001)
9. Garcia, J.N., Kutalik, Z., Cho, K.H., et al.: Level sets and minimum volume sets of probability density functions. International Journal of Approximate Reasoning 34(1), 25–47 (2003)
10. Pap, E., Surla, D.: Lebesgue measure of α-cuts approach for finding the height of the membership function. Fuzzy Sets and Systems 111(3), 341–350 (2000)
11. Lai, Y.J., Hwang, C.L.: Fuzzy Mathematical Programming-Methods and Applications. Springer, Berlin (1992)
12. Xu, Z.S.: Uncertain Multiple Attribute Decision Making: Methods and Applications. Tsinghua University Press, Beijing (2004) (in Chinese)
13. Dubois, D., Hüllermeier, E., Prade, H.: On the representation of fuzzy rules in terms of crisp rules. Information Sciences 151, 301–326 (2003)

14. Luo, C.Z., Wang, Z.P.: Representation of compositional relations in fuzzy reasoning. Fuzzy Sets and Systems 36(1), 77–81 (1990)
15. Wang, G.J.: Non-classical Logic and Approximate Reasoning. Science Press, Beijing (2000) (in Chinese)
16. Luo, C.Z.: Introduction to Fuzzy Sets (1). Beijing Normal University Press (1989)
17. Yuan, X.H., Li, H.X., Lee, E.S.: Three new cut sets of fuzzy sets and new theories of fuzzy sets. Computers and Mathematics with Applications 57, 691–701 (2009)
18. Yuan, X.H., Li, H.X., Sun, K.B.: Theory based on interval-valued level cut sets of Zadeh fuzzy sets. Advances in Intelligent and Soft Computing, vol. 62, pp. 501–510. Springer, Heidelberg (2009)
19. Goguen, J.A.: *L*-fuzzy Sets. J. Math. Anal. Appl. 18, 145–174 (1967)
20. Zadeh, L.A.: Outline of a new approach to the analysis of complex systems and decesion processes interval-valued fuzzy sets. IEEE Trans. Syst. Man Cybernet. 3, 28–44 (1973)
21. Atanassov, K.: Intuitionistic fuzzy sets. Fuzzy Sets and Systems 20, 87–96 (1986)
22. Atanassov, K.: Interval-valued intuitionistic fuzzy sets. Fuzzy Sets and Systems 31, 343–349 (1989)
23. Li, X.S., Yuan, X.H., Lee, E.S.: The three-dimensional fuzzy sets and their cut sets. Computers and Mathematics with Applications 58, 1349–1359 (2009)
24. Shang, Y.G., Yuan, X.H., Lee, E.S.: The n-dimensional fuzzy sets and Zadeh fuzzy sets based on the finite valued fuzzy sets. Computer and Mathematics with Applications (2010), doi:10,1016/j.camwa,2010,04,044
25. Mendel, J.M.: Advances in type-2 fuzzy sets and systems. Information Sciences 177, 84–110 (2007)
26. Yuan, X.H., Li, H.X., Sun, K.B.: The cut sets, decomposition theorems and representation theorems on intuitionistic fuzzy sets and interval-valued fuzzy Sets. Science in China, Series F 39(9), 933–945 (2009) (in Chinese)
27. Yuan, X.H., Li, H.X.: Cut sets on interval-valued intuitionistic fuzzy sets. In: Sixth International Conference On Fuzzy Systems and Knowledge Disvovery, vol. 6, pp. 167–171. IEEE Computer Society, Los Alamitos (2009)
28. Yuan, X.H., Li, H.X., Lee, E.S.: On the definition of intuitionistic subgroups. Computers and Mathematics with Applications 59(9), 3117–3129 (2010)

Approximation of Fuzzy Neural Networks to Fuzzy-Valued Measurable Function

Jun Li, Shuai Gao, and Bin Chen

School of Science, Communication University of China, Beijing 100024, China
e-mail: lijun@cuc.edu.cn

Abstract. In this note, we shall further investigate approximation property of regular fuzzy neural network (RFNN). It is shown that any fuzzy-valued measurable function can be approximated by four-layer RFNN in the sense of Sugeno integral norm for finite weakly null-additive fuzzy measure on $\mathbb{R}$. The previous result obtained by Li et al is improved.

Keywords: Fuzzy measure, Lusin's theorem, Approximation, Regular fuzzy neural network.

1 Introduction

In neural network theory, the learning ability of a neural network is closely related to its approximating capabilities, so it is important and interesting to study the approximation properties of neural networks. The studies on this matter were undertaken by many authors and a great number of important results were obtained ([1, 4, 12] etc). The similar approximation problems in fuzzy environment were investigated by Buckley [2, 3], P. Liu [8, 9] and other authors. In [9] Liu proved that continuous fuzzy-valued function can be closely approximated by a class of regular fuzzy neural networks (RFNNs) with real input and fuzzy-valued output. In this note, by using Lusin's theorem on fuzzy measure space, we show that such RFNNs is pan-approximator for fuzzy-valued measurable function. That is, any fuzzy-valued measurable function can be approximated by the four-layer RFNNs in the sense of Sugeno integral norm for the finite weakly null-additive measure on $\mathbb{R}$. The previous result we obtained in [7] is improved.

2 Preliminaries

We suppose that (X, ρ) is a metric space, and that $\mathcal{O}$ and $\mathcal{C}$ are the classes of all open and closed sets in (X, ρ), respectively, and $\mathcal{B}$ is Borel σ-algebra on X, i.e., it is the smallest σ-algebra containing $\mathcal{O}$.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 351–358.
springerlink.com

A set function $\mu : \mathcal{B} \to [0, +\infty]$ is called *weakly null-additive* ([13]), if $\mu(E \cup F) = 0$ whenever $E, F \in \mathcal{B}$ and $\mu(E) = \mu(F) = 0$; *subadditive* ([11]), if for any $E, F \in \mathcal{B}$ we have $\mu(E \cup F) \leq \mu(E) + \mu(F)$.

Obviously, the subadditivity of μ implies weak null-additivity.

A set function $\mu : \mathcal{B} \to [0, +\infty]$ is called *a fuzzy measure*, if it satisfies the following properties:

(FM1) $\mu(\emptyset) = 0$;

(FM2) $A \subset B$ implies $\mu(A) \leq \mu(B)$ (monotonicity);

(FM3) $A_1 \subset A_2 \subset \cdots$ implies $\lim_{n\to\infty} \mu(A_n) = \mu\left(\bigcup_{n=1}^{\infty} A_n\right)$ (continuity from below);

(FM4) $A_1 \supset A_2 \supset \cdots$, and there exists n_0 with $\mu(A_{n_0}) < +\infty$ imply

$$\lim_{n\to\infty} \mu(A_n) = \mu\left(\bigcap_{n=1}^{\infty} A_n\right) \text{ (continuity from above).}$$

When μ is a fuzzy measure, the triple $(X, \mathcal{B}, \mu)$ is called *a fuzzy measure space.*

In this paper, we always assume that μ is a finite fuzzy measure on $\mathcal{B}$.

Consider a nonnegative real-valued measurable function f on A and the *Sugeno integral* of f on A with respect to μ, which is denoted by

$$(S)\int_A f\, d\mu \triangleq \bigvee_{0\leq\alpha<+\infty} [\alpha \wedge \mu(A \cap F_\alpha)],$$

where $F_\alpha = \{x : f(x) \geq \alpha\}$.

Lemma 2.1 ([6]). *μ is weakly null-additive if and only if for any $\epsilon > 0$ and any double sequence $\{A_n^{(k)} \mid n \geq 1, k \geq 1\} \subset \mathcal{B}$ satisfying $A_n^{(k)} \searrow D_n$ $(k \to \infty)$, $\mu(D_n) = 0$, $n = 1, 2, \ldots$, there exists a subsequence $\{A_n^{(k_n)}\}$ of $\{A_n^{(k)} \mid n \geq 1, k \geq 1\}$ such that*

$$\mu\left(\bigcup_{n=1}^{\infty} A_n^{(k_n)}\right) < \epsilon \quad (k_1 < k_2 < \ldots).$$

Theorem 2.1 (Lusin's theorem [6]). *Let (X, ρ) be metric space and μ be weakly null-additive fuzzy measure on $\mathcal{B}$. If f is a real-valued measurable function on $E \in \mathcal{B}$, then, for every $\epsilon > 0$, there exists a closed subset $F_\epsilon \in \mathcal{B}$ such that f is continuous on F_ϵ and $\mu(E - F_\epsilon) < \epsilon$.*

3 Approximation in Fuzzy Mean by Regular Fuzzy Neural Networks

In this section, we study an approximation property of the four-layer RFNNs to fuzzy-valued measurable function in the sense of Sugeno integral norm for fuzzy measure on $\mathbb{R}$.

Let $\mathcal{F}_0(\mathbb{R})$ be the set of all bounded fuzzy numbers, i.e., for $\tilde{A} \in \mathcal{F}_0(\mathbb{R})$, the following conditions hold:

(i) $\forall\ \alpha \in (0,1]$, $\tilde{A}_\alpha \triangleq \{x \in \mathbb{R} \mid \tilde{A}(x) \geq \alpha\}$ is the closed interval of $\mathbb{R}$;
(ii) $\mathrm{Supp}(\tilde{A}) \triangleq \overline{\{x \in \mathbb{R} \mid \tilde{A}(x) > 0\}} \subset \mathbb{R}$ is a bounded set;
(iii) $\{x \in \mathbb{R} \mid \tilde{A}(x) = 1\} \neq \emptyset$.

For simplicity, $\mathrm{supp}(\tilde{A})$ is also written as $\tilde{A}_0$. Obviously, $\tilde{A}_0$ is a bounded and closed interval of $\mathbb{R}$. For $\tilde{A} \in \mathcal{F}_0(\mathbb{R})$, let $\tilde{A}_\alpha = [a_\alpha^-, a_\alpha^+]$ for each $\alpha \in [0,1]$ and we denote

$$|\tilde{A}| \triangleq \bigvee_{\alpha \in [0,1]} (|a_\alpha^-| \vee |a_\alpha^+|).$$

Proposition 3.1 ([9]). *Assume $\tilde{A}, \tilde{A}_1, \tilde{A}_2 \in \mathcal{F}_0(\mathbb{R})$, and $\tilde{W}_i, \tilde{V}_i \in \mathcal{F}_0(\mathbb{R})(i = 1, 2, \cdots, n)$. Then*
(1) $D(\tilde{A} \cdot \tilde{A}_1, \tilde{A} \cdot \tilde{A}_2) \leq |\tilde{A}| \cdot D(\tilde{A}_1, \tilde{A}_2)$,
(2) $D(\sum_{i=1}^n \tilde{W}_i, \sum_{i=1}^n \tilde{V}_i) \leq \sum_{i=1}^n D(\tilde{W}_i, \tilde{V}_i)$.

For $\tilde{A}, \tilde{B} \in \mathcal{F}_0(\mathbb{R})$, define metric $d(\tilde{A}, \tilde{B})$ between $\tilde{A}$ and $\tilde{B}$ by

$$d(\tilde{A}, \tilde{B}) \triangleq \sup_{\alpha \in [0,1]} d_H(\tilde{A}_\alpha, \tilde{B}_\alpha),$$

where d_H means Hausdorff metric: for $A, B \subset \mathbb{R}$,

$$d_H(A, B) \triangleq \max\left\{\sup_{x \in A} \inf_{y \in B} (|x - y|), \quad \sup_{y \in B} \inf_{x \in A} (|x - y|)\right\}.$$

It is known that $(\mathcal{F}_0(\mathbb{R}), d)$ is a completely separable metric space ([5]).

Let T be a measurable set in $\mathbb{R}$, $(T, \mathcal{B} \cap T, \mu)$ finite fuzzy measure space with weak null-additivity. Let $\mathcal{L}(T)$ denote the set of all fuzzy-valued measurable function

$$\tilde{F} : T \to \mathcal{F}_0(\mathbb{R}).$$

For any $\tilde{F}_1, \tilde{F}_2 \in \mathcal{L}(T)$, $d(\tilde{F}_1, \tilde{F}_2)$ is measurable function on $(T, \mathcal{B} \cap T)$, we will write a fuzzy integral norm as

$$\triangle_S(\tilde{F}_1, \tilde{F}_2) \triangleq (S) \int_T d(\tilde{F}_1, \tilde{F}_2) d\mu.$$

Proposition 3.2. *Let $\tilde{F}_1, \tilde{F}_2, \tilde{F}_3 \in \mathcal{L}(T)$. Then*

$$\triangle_S(\tilde{F}_1, \tilde{F}_3) \leq 2(\triangle_S(\tilde{F}_1, \tilde{F}_2) + \triangle_S(\tilde{F}_2, \tilde{F}_3)).$$

Proof. It is similar to the proof of Proposition 3.2 in [7].

Definition 3.1 ([9]). *A fuzzy-valued function $\tilde{\Phi} : T \to \mathcal{F}_0(\mathbb{R})$ is called a fuzzy-valued simple function, if there exist $\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_m \in \mathcal{F}_0(\mathbb{R})$, such that $\forall\ x \in T$,*

$$\tilde{\Phi}(x) = \sum_{k=1}^{m} \chi_{T_k}(x) \cdot \tilde{A}_k,$$

where $T_k \in \mathcal{B} \cap T$ $(k = 1, 2, \ldots, m)$, $T_i \cap T_j = \emptyset$ $(i \neq j)$ *and* $T = \bigcup_{k=1}^{m} T_k$, $\chi_{T_k}(x)$ *is the characteristic function of the set* T_k.

Denotes $\mathcal{S}(T)$ the set of all fuzzy-valued simple functions, then $\mathcal{S}(T) \subset \mathcal{L}(T)$.

Proposition 3.3. *Let* μ *be a finite, weakly null-additive fuzzy measure on* $\mathbb{R}^n$. *If* $\tilde{F} \in \mathcal{L}(T)$, *then for every* $\epsilon > 0$, *there exists* $\tilde{\Phi}_\epsilon \in \mathcal{S}(T)$ *such that*

$$\triangle_S(\tilde{F}, \tilde{\Phi}_\epsilon) < \epsilon.$$

Proof. By using weak null-additivity of μ, it is similar to the proof of Proposition 3.2.

Define

$$\mathcal{H}[\sigma] \triangleq \left\{ \tilde{H} \;\middle|\; \tilde{H}(x) = \sum_{i=1}^{n} \tilde{W}_i \cdot \left(\sum_{j=1}^{m} \tilde{V}_{ij} \cdot \sigma(x \cdot \tilde{U}_j + \tilde{\Theta}_j) \right) \right\},$$

where σ is a given extended function of $\sigma : \mathbb{R} \to \mathbb{R}$ (bounded, continuous and nonconstant), and $x \in \mathbb{R}, \tilde{W}_i, \tilde{V}_{ij}, \tilde{U}_j, \tilde{\Theta}_j \in \mathcal{F}_0(\mathbb{R})$.

For any $\tilde{H} \in \mathcal{H}[\sigma]$, $\tilde{H}$ is a four-layer feedforward RFNN with activation function σ, threshold vector $(\tilde{\Theta}_1, \ldots, \tilde{\Theta}_m)$ in the first hidden layer(cf. [9]).

Restricting fuzzy numbers $\tilde{V}_{ij}, \tilde{U}_j, \tilde{\Theta}_j \in \mathcal{F}_0(\mathbb{R})$, respectively, to be real numbers $v_{ij}, u_j, \theta_j \in \mathbb{R}$, we obtain the subset $\mathcal{H}_0[\sigma]$ of $\mathcal{H}[\sigma]$:

$$\mathcal{H}_0[\sigma] \triangleq \left\{ \tilde{H} \;\middle|\; \tilde{H}(x) = \sum_{i=1}^{n} \tilde{W}_i \cdot \left(\sum_{j=1}^{m} v_{ij} \cdot \sigma(x \cdot u_j + \theta_j) \right) \right\}.$$

Definition 3.2

(1) $\mathcal{H}_0[\sigma]$ *is call the pan-approximator of* $\mathcal{S}(T)$ *in the sense of* $\triangle_S$, *if* $\forall\ \tilde{\Phi} \in \mathcal{S}(T)$, $\forall\ \epsilon > 0$, *there exists* $\tilde{H}_\epsilon \in \mathcal{H}_0[\sigma]$ *such that*

$$\triangle_S(\tilde{\Phi}, \tilde{H}_\epsilon) < \epsilon.$$

(2) *For* $\tilde{F} \in \mathcal{L}(T)$, $\mathcal{H}[\sigma]$ *is call the pan-approximator for* $\tilde{F}$ *in the sense of* $\triangle_S$, *if* $\forall\ \epsilon > 0$, *there exists* $\tilde{H}_\epsilon \in \mathcal{H}[\sigma]$ *such that*

$$\triangle_S(\tilde{F}, \tilde{H}_\epsilon) < \epsilon.$$

By using Lusin's theorem (Theorem 2.1), Proposition 3.2 and 3.3 we can obtain the main result in this paper, which is stated in the following.

Theorem 3.1. *Let $(T, \mathcal{B} \cap T, \mu)$ be fuzzy measure space and μ be finite weakly null-additive fuzzy measure. Then,*

(1) $\mathcal{H}_0[\sigma]$ *is the pan-approximator of $\mathcal{S}(T)$ in the sense of $\triangle_S$.*
(2) $\mathcal{H}[\sigma]$ *is the pan-approximator for $\tilde{F}$ in the sense of $\triangle_S$.*

Proof. By using the conclusion of (1) and Proposition 3.3 we can obtain (2). Now we only prove (1).

Suppose that $\tilde{\Phi}(x)$ is a fuzzy-valued simple function, i.e.,

$$\tilde{\Phi}(x) = \sum_{k=1}^{m} \chi_{T_k}(x) \cdot \tilde{A}_k \quad (x \in T).$$

For arbitrarily given $\epsilon > 0$, applying Theorem 2.1 (Lusin's theorem) to each real measurable function $\chi_{T_k}(x)$, for every fixed k $(1 \le k \le m)$, there exists closed set $F_k \in \mathcal{B} \cap T$ such that

$$F_k \subset L_k \quad \text{and} \quad \mu(L_k - F_k) < \frac{\epsilon}{2m}$$

and $\chi_{T_k}(x)$ is continuous on F_k. Therefore, for every k there exist a Tauber-Wiener function σ and $p_k \in N, v'_{k1}, v'_{k2}, \cdots, v'_{kp_k}, \theta'_{k1}, \theta'_{k2}, \cdots, \theta'_{kp_k} \in \mathbb{R}$, and $\mathbf{w}'_{k1}, \mathbf{w}'_{k2}, \cdots, \mathbf{w}'_{kp_k} \in \mathbb{R}^n$ such that

$$\left| \chi_{T_k}(x) - \sum_{j=1}^{p_k} v'_{kj} \cdot \sigma(\langle \mathbf{w}'_{kj}, x \rangle + \theta'_{kj}) \right| < \frac{\epsilon}{2 \sum_{k=1}^{m} |\tilde{A}_k|} \quad (x \in L_k).$$

(We can assume $\sum_{k=1}^{m} |\tilde{A}_k| \neq 0$, without any loss of generality). Denote $L = \bigcap_{k=1}^{m} L_k$, then $T = L \cup (T - L)$. By Lemma 2.1, we have

$$\mu(T - L) = \mu\Big(\bigcup_{k=1}^{m} (T - L_k)\Big) \le \sum_{k=1}^{m} \mu(T - L_k) < \frac{\epsilon}{2}.$$

We take $\beta_1 = 0$, $\beta_k = \sum_{i=1}^{k-1} p_i, k = 2, \cdots, m$, and $p = \sum_{k=1}^{m} p_k$. For $k = 1, 2, \cdots, m, j = 1, 2, \cdots, p$, we denote

$$v_{kj} = \begin{cases} v'_{k(j-\beta_k)}, & \text{if } \beta_k < j \le \beta_{k+1}, \\ 0 & \text{otherwise,} \end{cases}$$

$$\theta_{kj} = \begin{cases} \theta'_{k(j-\beta_k)}, & \text{if } \beta_k < j \le \beta_{k+1}, \\ 0 & \text{otherwise,} \end{cases}$$

$$\mathbf{w}_{kj} = \begin{cases} \mathbf{w}'_{k(j-\beta_k)}, & \text{if } \beta_k < j \le \beta_{k+1}, \\ 0 & \text{otherwise,} \end{cases}$$

then, for any $k \in \{1, 2, \cdots, m\}$, we have

$$\sum_{j=1}^{p} v_{ij} \cdot \sigma(\langle \mathbf{w}_{kj}, x\rangle + \theta_{kj}) = \sum_{j=1}^{p_k} v_{ij}^{'} \cdot \sigma(\langle \mathbf{w}_{kj}^{'}, x\rangle + \theta_{kj}^{'}).$$

Define

$$\tilde{H}(x) = \sum_{k=1}^{m} \tilde{A}_k \cdot \left(\sum_{j=1}^{p} v_{kj} \cdot \sigma(\langle \mathbf{w}_{kj}, x\rangle + \theta_{kj}) \right),$$

then $\tilde{H} \in \mathcal{H}_0[\sigma]$.

In the following we prove $\triangle_S(\tilde{H}, \tilde{\Phi}) < \epsilon$.

Denote $B_{kj} = v_{kj} \cdot \sigma(\langle \mathbf{w}_{kj}, x\rangle + \theta_{kj})$ and $B_{kj}^{'} = v_{ij}^{'} \cdot \sigma(\langle \mathbf{w}_{kj}^{'}, x\rangle + \theta_{kj}^{'})$. By using Proposition 3.1 and noting $\mu(T - L) < \epsilon/2$, we have

$$\triangle_S(\tilde{H}, \tilde{\Phi}) = (S)\int_T d(\tilde{H}, \tilde{\Phi}) d\mu = \bigvee_{0 \le \alpha < +\infty} \left[\alpha \wedge \mu(T \cap (d(\tilde{H}, \tilde{\Phi}))_\alpha) \right]$$

$$= \bigvee_{0 \le \alpha < +\infty} \left[\alpha \wedge \mu \left(T \cap \left(d(\sum_{k=1}^{m} \tilde{A}_k \cdot \sum_{j=1}^{p} B_{kj}, \sum_{k=1}^{m} \chi_{T_k}(x) \cdot \tilde{A}_k) \right)_\alpha \right) \right]$$

$$\le \bigvee_{0 \le \alpha < +\infty} \left[\alpha \wedge \mu \left((L \cup (T - L)) \cap \left(\sum_{k=1}^{m} |\tilde{A}_k| \cdot d(\sum_{j=1}^{p} B_{kj}, \chi_{T_k}(x)) \right)_\alpha \right) \right]$$

$$\le \bigvee_{0 \le \alpha < +\infty} \left[\alpha \wedge \mu \left(L \cap \left(\sum_{k=1}^{m} |\tilde{A}_k| \cdot d(\sum_{k=1}^{p} B_{kj}, \chi_{T_k}(x)) \right)_\alpha \right) \right]$$

$$+ \bigvee_{0 \le \alpha < +\infty} \left[\alpha \wedge \mu \left((T - L) \cap \left(\sum_{k=1}^{m} |\tilde{A}_k| \cdot d(\sum_{k=1}^{p} B_{kj}, \chi_{T_k}(x)) \right)_\alpha \right) \right]$$

$$\le \bigvee_{0 \le \alpha < +\infty} \left[\alpha \wedge \mu \left(L \cap \left(\sum_{k=1}^{m} |\tilde{A}_k| \cdot |\sum_{j=1}^{p} B_{kj} - \chi_{T_k}(x)| \right)_\alpha \right) \right]$$

$$+ \bigvee_{0 \le \alpha < +\infty} \left[\alpha \wedge \mu \left((T - L) \right) \right]$$

$$\le \bigvee_{0 \le \alpha < +\infty} \left[\alpha \wedge \mu \left(L \cap \left(\sum_{k=1}^{m} |\tilde{A}_k| \cdot |\sum_{j=1}^{p_k} B_{kj}^{'} - \chi_{T_k}(x)| \right)_\alpha \right) \right] + \frac{\epsilon}{2}.$$

Now we estimate the first part in the above formula. If $x \in L$, then for every $k = 1, 2, \cdots, m$, we have $x \in L_k$, hence

$$\left| \chi_{T_k}(x) - \sum_{j=1}^{p_k} v_{kj}^{'} \cdot \sigma(\langle \mathbf{w}_{kj}^{'}, x\rangle + \theta_{kj}^{'}) \right| < \frac{\epsilon}{2\sum_{k=1}^{m} |\tilde{A}_k|},$$

for every $k = 1, 2, \cdots, m$. That is, for $x \in L$,

$$\sum_{k=1}^{m} |\tilde{A}_k| \cdot |\sum_{j=1}^{p_k} B_{kj}^{'} - \chi_{T_k}(x)| < \frac{\epsilon}{2}.$$

Therefore,

$$\begin{aligned}
&\bigvee_{0\le\alpha<+\infty} \left[\alpha \wedge \mu\left(L \cap \left(\sum_{k=1}^{m} |\tilde{A}_k| \cdot |\sum_{j=1}^{p_k} B_{kj}^{'} - \chi_{T_k}(x)|\right)_{\alpha}\right)\right] \\
&= \bigvee_{\alpha\in[0,\frac{\epsilon}{2}]} \left[\alpha \wedge \mu\left(L \cap \left(\sum_{k=1}^{m} |\tilde{A}_k| \cdot |\sum_{j=1}^{p_k} B_{kj}^{'} - \chi_{T_k}(x)|\right)_{\alpha}\right)\right] \\
&\quad + \bigvee_{\alpha\in[\frac{\epsilon}{2},\infty)} \left[\alpha \wedge \mu\left(L \cap \left(\sum_{k=1}^{m} |\tilde{A}_k| \cdot |\sum_{j=1}^{p_k} B_{kj}^{'} - \chi_{T_k}(x)|\right)_{\alpha}\right)\right] \\
&= \bigvee_{\alpha\in[0,\frac{\epsilon}{2}]} \left[\alpha \wedge \mu\left(L \cap \left(\sum_{k=1}^{m} |\tilde{A}_k| \cdot |\sum_{j=1}^{p_k} B_{kj}^{'} - \chi_{T_k}(x)|\right)_{\alpha}\right)\right] \\
&\le \frac{\epsilon}{2}.
\end{aligned}$$

Thus, we obtain $\triangle_S(\tilde{H}, \tilde{\Phi}) < \epsilon$. The proof of (1) now is complete.

Acknowledgments. This work was partially supported by the NSFC Grant No.70771010.

References

1. Attali, J.G., Pages, G.: Approximation of functions by a multilayer perceptron: a new approach. Neural Networks 10, 1069–1081 (1997)
2. Buckley, J.J., Hayashi, Y.: Fuzzy input-output controllers are universal approximators. Fuzzy Sets and Systems 58, 273–278 (1993)
3. Buckley, J.J., Hayashi, Y.: Can fuzzy neural nets approximate continuous fuzzy function. Fuzzy Sets and Systems 61, 43–51 (1994)
4. Burton, R.M., Dehling, H.G.: Universal approximation in p-mean by neural networks. Neural Networks 11, 661–667 (1998)
5. Diamond, P., Kloeden, P.: Metric Spaces of Fuzzy Sets. World Scientific Publishing, Singapore (1994)

6. Li, J., Yasuda, M.: Lusin's theorem on fuzzy measure spaces. Fuzzy Sets and Systems 146, 121–133 (2004)
7. Li, J., Li, J.Z., Yasuda, M.: Approximation of fuzzy neural networks using Lusin's theorem. In: Proceedings of IPMU 2006 International Conference, EDK, Paris, pp. 2485–2490 (2006)
8. Liu, P.: Analyses of regular fuzzy neural networks for approximation capabilities. Fuzzy Sets and Systems 114, 329–338 (2000)
9. Liu, P.: Universal approximations of continuous fuzzy-valued functions by multi-layer regular fuzzy neural networks. Fuzzy Sets and Systems 119, 313–320 (2001)
10. Liu, P.: Approximation analyses of feedforward regular fuzzy neural network with two hidden layers. Fuzzy Sets and Systems 150, 373–396 (2005)
11. Pap, E.: Null-additive Set Functions. Kluwer, Dordrecht (1995)
12. Scarselli, F., Tsoi, A.G.: Universal approximation using feedforward neural networks: a survey of some existing methods, and some new results. Neural Networks 11, 15–17 (1998)
13. Wang, Z.Y., Klir, G.J.: Fuzzy Measure Theory. Plenum, New York (1992)

Fuzzy Divergences Based on Tsallis Relative Entropy

Lan Rong and Fan Jiu-lun

School of Communication and Information Engineering, Xi'an Institute of Post and Telecommunications, Xi'an 710121, China
e-mail: ronglanlogic@163.com, jiulunf@xupt.edu.cn

Abstract. Fuzzy divergence describes the difference between two fuzzy sets. Based on Tsallis relative entropy, two classes of fuzzy divergence are proposed in this paper. Using the point of view of f-divergence proposed by Csiszar, the properties of the two classes of fuzzy divergence are discussed. The notable feature of the two classes of fuzzy divergence is that they are both σ-distance measures, thus new fuzzy entropy formulas can be induced by these fuzzy divergences.

Keywords: Fuzzy divergence; Tsallis relative entropy; Distance measure; Fuzzy entropy.

1 Introduction

In order to measure the difference between two fuzzy sets, Bhandari et al [1,2] introduced fuzzy divergence using logarithm operation, they also gave an application to image segmentation. Another fuzzy divergence using exponential operation was proposed by us [3] and an application to image segmentation was given by Charia and Ray [4].

Within the framework of multifractal, the quantity that is normally scaled is p_i^q, where p_i is the probability associated with an event and $q \in \mathbb{R}$. Tsallis used this quantity to generalize the Shannon entropy in information theory, and presented the Tsallis entropy[5]. Since then, Tsallis entropy has been widely studied and applied in a variety of substantive areas. In this paper, two new classes of divergence are proposed to measure the difference between two fuzzy sets as the extension of the Tsallis entropy in fuzzy cases. In view of the f-divergence proposed by Csiszar [6], we study properties of the presented fuzzy divergences which take χ^2-divergences as special cases. As an application, we state some ways for the fuzzy entropies [7] generated by the two Tsallis fuzzy divergences.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 359–368.
springerlink.com

2 Entropy and Distance Measure

In this section, we first introduce the concepts of the fuzzy set theory. Throughout this paper, $\mathbb{R}_+ = [0, \infty)$.

Definition 1. *Let X be an universe, a fuzzy set A defined on X is given by $A = \{(x, \mu_A(x)) | x \in X\}$, where the mapping $\mu_A : X \longrightarrow [0, 1]$ is called the membership function. $A^c$is the complement of A, i.e. $\mu_{A_c}(x) = 1 - \mu_A(x)$, $\forall x \in X$.*

Now let $F(X)$ denote the set of all fuzzy subsets on X. Let $P(X)$ denote the set of all crisp sets on X. $[a]$ is the fuzzy set on X for which $\mu_{[a]}(x) = a$, $\forall x \in X (0 \leq a \leq 1)$.

Definition 2. *A fuzzy set A^* is called a sharpening of A, if $\mu_{A^*}(x) \geq \mu_A(x)$ when $\mu_A(x) \geq \frac{1}{2}$, and $\mu_{A^*}(x) \leq \mu_A(x)$ when $\mu_A(x) < \frac{1}{2}$.*

Definition 3. *$\forall A \in F(X)$, crisp sets $A_{near}, A_{far} \in P(X)$, are defined as*

$$\mu_{A_{near}}(x) = \begin{cases} 1, & \mu_A(x) \geq \frac{1}{2}, \\ 0, & \mu_A(x) < \frac{1}{2}, \end{cases} \mu_{A_{far}}(x) = \begin{cases} 0, & \mu_A(x) \geq \frac{1}{2}, \\ 1, & \mu_A(x) < \frac{1}{2}. \end{cases}$$

In fuzzy set theory, fuzzy entropy [7] is a very basic concept to measure the fuzziness of a fuzzy set.

Definition 4. *The mapping $e : F(X) \longrightarrow \mathbb{R}_+$ is called an entropy on $F(X)$ if e has the following properties:*
1) $e(G) = 0, \forall G \in P(X)$;
2) $e([\frac{1}{2}]) = \max_{A \in F(X)} e(A)$;
3) $e(A^*) \leq e(A)$ for any sharpening A^* of A;
4) $e(A) \leq e(A^c)$, $\forall A \in F(X)$.

Definition 5. *Let e be an entropy on $F(X)$. e is called a $\sigma-$entropy on $F(X)$ if $e(A) = e(A \cap G) + e(A \cap G^c)$, $\forall G \in P(X)$.*

Definition 6. *The mapping $d : F(X) \times F(X) \longrightarrow \mathbb{R}_+$ is called a distance measure on $F(X)$ if d has the following properties:*
1) $d(A, B) = d(B, A), \forall A, B \in F(X)$;
2) $d(A, A) = 0, \forall A \in F(X)$;
3) $d(G, G^c) = \max_{A,B \in F(X)} d(A, B), \forall G \in P(X)$;
4) If $A \subseteq B \subseteq C$, then $d(A, B) \leq d(A, C)$ and $d(B, C) \leq d(A, C)$.

Definition 7. *Let d be a distance measure on $F(X)$. d is called a $\sigma-$distance measure on $F(X)$ if $\forall A, B \in F(X)$ and $\forall G \in P(X)$, we have*

$$d(A, B) = d(A \cap G, B \cap G) + d(A \cap G^c, B \cap G^c).$$

In [7] and [9], it had shown that $d(A, B) = d(A \cap B, A \cup B)$, $d(A, B) + d(A, C) = d(A, B \cap C) + d(A, B \cup C)$ if is a $\sigma-$distance measure.

Let e be an entropy, e is called a normalized entropy if $0 \le e(A) \le 1$. Let d be a distance measure, d is called a normalized distance measure if $0 \le d(A, B) \le 1$. $\forall G \in P(X)$, $\frac{1}{2}G$ is defined by $\mu_{\frac{1}{2}G}(x) = \begin{cases} \frac{1}{2}, & x \in G \\ 0, & x \in G^c \end{cases}$. We had proved:

Theorem 1. *Let d be a $\sigma-$distance measure on $F(X)$. If d satisfies the following properties*
1)$d(\frac{1}{2}G, [0]) = d(\frac{1}{2}G, G)$, $\forall G \in P(X)$;
2)$d(A^c, B^c) = d(A, B)$,$\forall A, B \in F(X)$,
then $e(A) = \frac{d(A, A_{near})}{d(A, A_{far})}$ and $e(A) = d(A, A_{near}) + 1 - d(A, A_{far})$ are normalized entropies.

3 Fuzzy Divergence Based on Tsallis Relative Entropy

In this section, two classes of fuzzy divergence based on Tsallis entropy form are proposed. The properties of the proposed fuzzy divergences are discussed. And then we research the relationships between fuzzy divergence and entropy, distance measure.

In the field of the statistical physics, the Tsallis entropy [10] is defined as following:

$$S_q(X) = \Sigma_x p(x)^q \ln_q p(x),$$

where $ln_q p(x) \equiv (x^{1-q} - 1)/(1 - q)$ for any non-negative real number q and x, and $p(x) \equiv p(X = x)$is the probability distribution of the given random variable X. We easily find that the Tsallis entropy converges to the Shannon entropy as $q \to 1$.

Let $\Delta_n = \{P = (p_1, p_2, \cdots, p_n) | p_i > 0, \Sigma_{i=1}^n p_i = 1\}$, $n \ge 2$ be the set of complete finite discrete probability distributions. $\forall P, Q \in \Delta_n$, Tsallis relative entropy is defined as:

$$K_q(P \parallel Q) = -\sum_{i=1}^{n} p_i \ln_q \frac{q_i}{p_i} = \frac{1}{q-1} \sum_{i=1}^{n} p_i((\frac{q_i}{p_i})^{1-q} - 1), q > 0, q \ne 1. \quad (1)$$

When $q = 2$, Eq. (1) is the χ^2-divergence:

$$\chi^2(P \parallel Q) = \sum_{i=1}^{n} \frac{(p_i - q_i)^2}{q_i} = K_2(P \parallel Q).$$

In the next section, we will introduce the Tsallis fuzzy divergences as the extension of the Tsallis relative entropy, and show their properties.

3.1 The Novel Tsallis Fuzzy Divergences

Let $X = \{x_1, x_2, \ldots, x_n\}$ be an universe and $A, B \in F(X)$. Then the fuzzy information in favor of A against B is defined as

$$K_{1q}(A,B) = \frac{1}{q-1}\sum_{i=1}^{n}\mu_A(x_i)[(\frac{\mu_A(x_i)}{\mu_B(x_i)})^{q-1} - 1], q > 0, q \neq 1.$$

So we can define the divergence between two fuzzy sets by

$$K_q(A,B) = K_{1q}(A,B) + K_{1q}(A^c,B^c).$$

Note that the above formula does not include the crisp set, in order to account for this, two modified expression for $K_q(A,B)$ are shown

$$K_q^{'}(A,B) = \frac{1}{q-1}\sum_{i=1}^{n}(\frac{(1+\mu_A(x_i))^q}{(1+\mu_B(x_i))^{q-1}} + \frac{(2-\mu_A(x_i))^q}{(2-\mu_B(x_i))^{q-1}} - 1),$$

$$K_q^{''}(A,B) = \frac{1}{q-1}\sum_{i=1}^{n}(\frac{2^{q-1}\mu_A^q(x_i)}{(\mu_A(x_i)+\mu_B(x_i))^{q-1}} + \frac{2^{q-1}(1-\mu_A(x_i))^q}{(2-\mu_A(x_i)-\mu_B(x_i))^{q-1}} - 1).$$

However, we observe that $K_q^{'}(A,B) \neq K_q^{'}(B,A)$ and $K_q^{''}(A,B) \neq K_q^{''}(B,A)$. Therefore, two symmetric measures are defined as follows

$$D_1(A,B) = K_q^{'}(A,B) + K_q^{'}(B,A), q > 0, q \neq 1, \quad (2)$$

$$D_2(A,B) = K_q^{''}(A,B) + K_q^{''}(B,A), q > 0, q \neq 1. \quad (3)$$

3.2 The Properties of the Tsallis Fuzzy Divergences

Now we discuss the properties of the $D_1(A,B)$ and $D_2(A,B)$. For this, we introduce the following lemma.

Given a convex function $f : [0,\infty) \to \mathbb{R}$, the $f-$divergence measure introduced by Csiszar [6] is given by

$$C_f(p,q) = \sum_{i=1}^{n} q_i f(\frac{p_i}{q_i}), p, q \in \mathbb{R}_+^n.$$

For $C_f(p,q)$, Csiszar and Korner [11] showed the following two lemmas.

Lemma 1. *(Joint convexity) Let* $f : [0,\infty) \to \mathbb{R}$ *be convex function. Then* $C_f(p,q)$ *is jointly convex in* p *and* q*, where* $p, q \in \mathbb{R}_+^n$.

Lemma 2. *(Jensen's inequality) Let* $f : [0,\infty) \to \mathbb{R}$ *be convex function,* $\forall p, q \in \mathbb{R}_+^n$ *with* $P_n = \Sigma_{i=1}^n p_i > 0$*,* $Q_n = \Sigma_{i=1}^n q_i > 0$*. We have the inequality*

$$C_f(p,q) \geq Q_n f(\frac{P_n}{Q_n}).$$

The equality sign holds if and only if

$$\frac{p_1}{q_1} = \frac{p_2}{q_2} = \ldots = \frac{p_n}{q_n}.$$

In particular, $\forall P, Q \in \Delta_n$, we have $C_f(P\|Q) \triangleq C_f(P,Q) \geq f(1)$, with equality if and only if $P = Q$.

In view of the above lemmas, we can state the properties of $D_1(A,B)$. Let
$p = (1+\mu_B(x_i), 2-\mu_B(x_i), 1+\mu_A(x_i), 2-\mu_A(x_i))$,
$q = (1+\mu_A(x_i), 2-\mu_A(x_i), 1+\mu_B(x_i), 2-\mu_B(x_i))$.
For x_i in $D_1(A,B)$, we have

$$\begin{aligned} D_1(A,B;x_i) &= \frac{1}{q-1}((\frac{(1+\mu_A(x_i))^q}{(1+\mu_B(x_i))^{q-1}} + \frac{(2-\mu_A(x_i))^q}{(2-\mu_B(x_i))^{q-1}} - 1) \\ &\quad +(\frac{(1+\mu_B(x_i))^q}{(1+\mu_A(x_i))^{q-1}} + \frac{(2-\mu_B(x_i))^q}{(2-\mu_A(x_i))^{q-1}} - 1)) \\ &= \sum_{j=1}^{4} q_j f(\frac{p_j}{q_j}), \end{aligned}$$

where the function $f_1 : [0,\infty) \to \mathbb{R}$ is defined by $f_1(t) = \frac{1}{q-1}(t^{1-q} - \frac{1}{3}), q > 0, q \neq 1$. On account of $f_1^{''}(t) = qt^{-q-1} \geq 0$ and **Lemma 1**, $D_1(A,B;x_i)$ is jointly convex in p and q. According to **Lemma 2**, we have

$$D_1(A,B;x_i) \geq Q_n f(\frac{P_n}{Q_n}) = 6f(1) = \frac{4}{q-1},$$

equality sign holds if and only if

$$\frac{p_1}{q_1} = \frac{p_2}{q_2} = \frac{p_3}{q_3} = \frac{p_4}{q_4}, i.e., \mu_A(x_i) = \mu_B(x_i).$$

The above conclusion shows that the minimum of $D_1(A,B;x_i)$ is $\frac{4}{\alpha-1}(\neq 0)$. In order to make the minimum of $D_1(A,B)$ be 0, a modified expression $D_1'(A,B)$ is shown

$$\begin{aligned} D_1'(A,B) &= \frac{1}{q-1}(\sum_{i=1}^{n}(\frac{(1+\mu_A(x_i))^q}{(1+\mu_B(x_i))^{q-1}} + \frac{(2-\mu_A(x_i))^q}{(2-\mu_B(x_i))^{q-1}} - 3) \\ &\quad + \sum_{i=1}^{n}(\frac{(1+\mu_B(x_i))^q}{(1+\mu_A(x_i))^{q-1}} + \frac{(2-\mu_B(x_i))^q}{(2-\mu_A(x_i))^{q-1}} - 3)), q > 0, q \neq 1. \end{aligned}$$

In the following we will give a complete proof for the conclusion that $D_1'(A,B)$ is a $\sigma-$distance measure, although part results had been stated in [1,2].

Lemma 3. *Let* $f : [0,1]^2 \to \mathbb{R}_+$ *satisfies* $\forall (x,y) \in [0,1]^2$, $f(x,y) = f(y,x)$ *and* $f(x,y) = f(1-x, 1-y)$, *then* $f(x,y) \leq f(x,z)$ *if and only if* $f(y,z) \leq f(x,z)$ *for* $0 \leq x \leq y \leq z \leq 1$.

For stating the properties of $D_1'(A,B;x_i)$, we present a function $f(x,y)$ as follows.

Lemma 4. *$\forall x, y \in [0,1]$, let*

$$f(x,y) = \frac{1}{q-1}\left(\frac{(1+x)^q}{(1+y)^{q-1}} + \frac{(1+y)^q}{(1+x)^{q-1}} + \frac{(2-x)^q}{(2-y)^{q-1}} + \frac{(2-y)^q}{(2-x)^{q-1}} - 6\right).$$

If $0 \leq x \leq y \leq z \leq 1$, then $f(x,y) \leq f(x,z)$ and $f(y,z) \leq f(x,z)$.

Proof. For f, we can get directly that $\forall (x,y) \in [0,1]^2$, $f(x,y) = f(y,x)$ and $f(x,y) = f(1-x, 1-y)$.

Suppose $x \leq y$, then

$$\frac{(2-x)}{(2-y)} \geq 1 \geq \frac{(1+x)}{(1+y)}, \frac{(2-y)}{(2-x)} \leq 1 \leq \frac{(1+y)}{(1+x)}.$$

We have

$$f_y(x,y) = [(\frac{2-x}{2-y})^q - (\frac{1+x}{1+y})^q] + \frac{q}{q-1}[(\frac{1+y}{1+x})^{q-1} - (\frac{2-y}{2-x})^{q-1}].$$

The function $x^q (x \geq 0)$ is a monotonically increasing function for $q > 0$ and $q \neq 1$, so we have

$$(\frac{2-x}{2-y})^q - (\frac{1+x}{1+y})^q \geq 0.$$

When $q > 1$, then $q - 1 > 0$ and $x^{q-1} (x \geq 0)$ is a monotonically increasing function. We have

$$\frac{q}{q-1}[(\frac{1+y}{1+x})^{q-1} - (\frac{2-y}{2-x})^{q-1}] \geq 0.$$

When $0 < q < 1$, then $q - 1 < 0$ and $x^{q-1} (x \geq 0)$ is a monotonically decreasing function. We have

$$\frac{q}{q-1}[(\frac{1+y}{1+x})^{q-1} - (\frac{2-y}{2-x})^{q-1}] \geq 0.$$

Therefore, $f(x,y)$ is a monotonically increasing function on y for $x \leq y$. So we have $f(x,y) \leq f(x,z)$ for $0 \leq x \leq y \leq z \leq 1$. According to **Lemma 3**, we can obtain $f(y,z) \leq f(x,z)$. Thus, the proof is complete.

Theorem 2. *$\forall A, B, C \in F(X)$, the divergence measure $D_1'(A,B)$ satisfies the following properties*

1) $D_1'(A,B) \geq 0$, $D_1'(A,B) = 0$ if and only if $A = B$;
2) $\forall G \in P(X)$, $D_1'(G, G^c) = \max_{A,B \in F(X)} D_1'(A,B) = \frac{2n}{q-1}(2^q + 2^{1-q} - 3)$;
3) $D_1'(A,B) = D_1'(B,A)$;
4) $D_1'(A,B) = D_1'(A^c, B^c)$;
5) If $A \subseteq B \subseteq C$, then $D_1'(A,B) \leq D_1'(A,C)$ and $D_1'(B,C) \leq D_1'(A,C)$;
6) $D_1'([\frac{1}{2}], [1]) = D_1'([\frac{1}{2}], [0])$.

Proof. According to **Lemma 1** and **Lemma 2**, we get 1). According to **Lemma 3** and **Lemma 4**, we get 2) and 5). Obviously, we obtain 3), 4) and 6).

According to **Theorem 2**, we know that $D_1'(A,B)$ is a distance measure, and we have the following conclusion.

Theorem 3. *$D_1'(A,B)$ is a $\sigma-$distance measure.*

Let
$$\overline{D}_1'(A,B) = \frac{q-1}{2n(2^q + 2^{1-q} - 3)} D_1'(A,B).$$

Then $\overline{D}_1'(A,B)$ is a normal distance measure. According to **Theorem 2**, we have

Proposition 1. *$\overline{D}_1'(A,B)$ has the following properties:*
1)$\overline{D}_1'([\frac{1}{2}]G,[1]) = \overline{D}_1'([\frac{1}{2}]G,[0])$, $\forall G \in P(X)$;
2)$\overline{D}_1'(A,B) = \overline{D}_1'(A^c,B^c)$, $\forall A,B \in F(X)$.

By **Proposition 1**, **Theorem 1** and **Theorem 2**, we get

$$e(A) = \frac{\overline{D}_1'(A,A_{near})}{\overline{D}_1'(A,A_{far})}$$

and
$$e(A) = \overline{D}_1'(A,A_{near}) + 1 - \overline{D}_1'(A,A_{far})$$

are fuzzy entropies.

For $D_2(A,B)$, we can define a function $h:[0,1]^2 \to \mathbb{R}_+$ by

$$\begin{aligned} h(x,y) = \frac{2^{q-1}}{q-1}[&\frac{x^q}{(x+y)^{q-1}} + \frac{(1-x)^q}{(2-x-y)^{q-1}} \\ &+\frac{y^q}{(x+y)^{q-1}} + \frac{(1-y)^q}{(2-x-y)^{q-1}} - 2^{2-q}]. \end{aligned}$$

We can also prove that $h(x,y)$ has the following property.

Lemma 5. *If $0 \le x \le y \le z \le 1$, then $h(x,y) \le h(x,z)$ and $h(y,z) \le h(x,z)$.*

Proof. Suppose $x \le y$, we have

$$\begin{aligned} h_y(x,y) = 2^{q-1}[(\frac{1-x}{2-x-y})^q - (\frac{x}{x+y})^q] + 2^{q-1}[(\frac{q}{q-1}(\frac{y}{x+y})^{q-1} \\ -(\frac{y}{x+y})^q) - (\frac{q}{q-1}(\frac{1-y}{2-x-y})^{q-1} - (\frac{1-y}{2-x-y})^q)]. \end{aligned}$$

By $x \leq y$, we get

$$\frac{1-x}{2-x-y} \geq \frac{x}{x+y}.$$

The function $x^q(x \geq 0)$is a monotonically increasing function for $q > 0$ and $q \neq 1$, we have

$$2^{q-1}[(\frac{1-x}{2-x-y})^q - (\frac{x}{x+y})^q] \geq 0.$$

Let $H(x) = \frac{q}{q-1}x^{q-1} - x^q$ for $x \in [0,1]$. According $H'(x) = qx^{q-2}(1-x) \geq 0$, we know that $H(x)$ is a monotonically increasing function. By

$$0 \leq \frac{1-x}{2-x-y} \leq \frac{x}{x+y} \leq 1,$$

we get

$$\frac{q}{q-1}(\frac{y}{x+y})^{q-1} - (\frac{y}{x+y})^q \geq \frac{q}{q-1}(\frac{1-y}{2-x-y})^{q-1} - (\frac{1-y}{2-x-y})^q.$$

So we have

$$2^{q-1}[(\frac{q}{q-1}(\frac{y}{x+y})^{q-1} - (\frac{y}{x+y})^q) - (\frac{q}{q-1}(\frac{1-y}{2-x-y})^{q-1} - (\frac{1-y}{2-x-y})^q)] \geq 0.$$

When $x \leq y$, we have $h_y(x,y) \geq 0$. So $h(x,y)$ is a monotonically increasing function on y, i.e., $h(x,y) \leq h(x,z)$ for $0 \leq x \leq y \leq z \leq 1$. By **Lemma 3**, we have $h(y,z) \leq h(x,z)$.

Theorem 4. *$\forall A, B, C \in F(X)$, the divergence measure $D_2(A,B)$ satisfies the following properties*

1) $D_2(A,B) \geq 0$, $D_2(A,B) = 0$ if and only if $A = B$;

2) $\forall G \in P(X)$, $D_2(G, G^c) = \max_{A,B \in F(X)} D_2(A,B) = \frac{2n}{q-1}(2^{1-q} - 1)$;

3) $D_2(A,B) = D_2(B,A)$;

4) $D_2(A,B) = D_2(A^c, B^c)$;

5) If $A \subseteq B \subseteq C$, then $D_2(A,B) \leq D_2(A,C)$ and $D_2(B,C) \leq D_2(A,C)$;

6) $D_2([\frac{1}{2}],[1]) = D_2([\frac{1}{2}],[0])$.

Proof. For x_i in $D_2(A,B)$, we have

$$\begin{aligned} D_2(A,B;x_i) = \frac{1}{q-1}((&\frac{2^{q-1}\mu_A^q(x_i)}{(\mu_A(x_i)+\mu_B(x_i))^{q-1}} + \frac{2^{q-1}(1-\mu_A(x_i))^q}{(2-\mu_A(x_i)-\mu_B(x_i))^{q-1}} - 1) \\ &+(\frac{2^{q-1}\mu_B^q(x_i)}{(\mu_A(x_i)+\mu_B(x_i))^{q-1}} + \frac{2^{q-1}(1-\mu_B(x_i))^q}{(2-\mu_B(x_i)-\mu_A(x_i))^{q-1}} - 1)). \end{aligned}$$

Given

$$p = (\frac{\mu_A(x_i)+\mu_B(x_i)}{2}, \frac{2-\mu_A(x_i)-\mu_B(x_i)}{2}), q = (\mu_A(x_i), 1-\mu_A(x_i)),$$

then we have

$$K_q''(A,B;x_i) = \frac{1}{q-1}(\frac{2^{q-1}\mu_A^q(x_i)}{(\mu_A(x_i)+\mu_B(x_i))^{q-1}} + \frac{2^{q-1}(1-\mu_A(x_i))^q}{(2-\mu_A(x_i)-\mu_B(x_i))^{q-1}} - 1)$$
$$= \sum_{j=1}^{2} q_j g_1(\frac{p_j}{q_j}),$$

where the function $g_1 : [0,\infty) \to \mathbb{R}$ is defined by $g_1(t) = \frac{1}{q-1}(t^{1-q}-1), q > 0$ and $q \neq 1$. According to $g_1^{''}(t) = qt^{-q-1} \geq 0$ and **Lemma 1**, we know that $K_q''(A,B;x_i)$ is jointly convex in p and q. By **Lemma 2**, we have $K_q''(A,B;x_i) \geq g(1) = 0$, equality sign holds if and only if

$$\frac{p_1}{q_1} = \frac{p_2}{q_2}, i.e., \mu_A(x_i) = \mu_B(x_i).$$

Given

$$p = (\frac{\mu_A(x_i)+\mu_B(x_i)}{2}, \frac{2-\mu_A(x_i)-\mu_B(x_i)}{2}), q = (\mu_B(x_i), 1-\mu_B(x_i)),$$

we know that $K_q''(B,A;x_i)$ is jointly convex in p and q, and $K_q''(B,A;x_i) \geq g_1(1) = 0$, equality sign holds if and only if $\mu_A(x_i) = \mu_B(x_i)$.

Therefore, $D_2(A,B;x_i) \geq 0$ for any $x_i \in X$, $D_2(A,B;x_i) = 0$ if and only if $\mu_A(x_i) = \mu_B(x_i)$. And we obtain 1).

According to **Lemma 5**, we have 2) and 5). Obviously, we get 3), 4) and 6).

According to **Theorem 4**, we know that $D_2(A,B)$ is a distance measure, and we have the following conclusion.

Theorem 5. $D_2(A,B)$ *is a* $\sigma-$*distance measure.*

Let

$$\overline{D}_2(A,B) = \frac{q-1}{2n(2^{1-q}-1)} D_2(A,B),$$

then $\overline{D}_2(A,B)$ is a normal distance measure. According to **Theorem 2**, we have

Proposition 2. $\overline{D}_2(A,B)$ *has the following properties:*

1)$\overline{D}_2([\frac{1}{2}]G,[1]) = \overline{D}_2([\frac{1}{2}]G,[0])$, $\forall G \in P(X)$;

2)$\overline{D}_2(A,B) = \overline{D}_2(A^c,B^c)$, $\forall A,B \in F(X)$.

Therefore

$$e(A) = \frac{\overline{D}_2(A,A_{near})}{\overline{D}_2(A,A_{far})}$$

and

$$e(A) = \overline{D}_2(A,A_{near}) + 1 - \overline{D}_2(A,A_{far})$$

are fuzzy entropies.

4 Conclusion

Distance measure is a basic concept in fuzzy set theory. Based on Tsallis relative entropy, two new classes of fuzzy divergence measure are proposed in this paper. The properties of the proposed divergence measures are discussed. we show that the proposed divergence measures are $\sigma-$distance measures. By the two $\sigma-$distance measures, we can induce some new fuzzy entropy. In this paper, we discuss the fuzzy divergence with parameter. Considering fuzzy divergences had been used in image segmentation, in the next work, we will apply the proposed distance measure, more clearly $\sigma-$distance measure, to image segmentation or other areas.

Acknowledgments. The work is supported by the Education Department Science Research Scheme Foundation of Shaanxi Province of China under Grant No. 09JK720.

References

1. Bhandari, D., et al.: Fuzzy deivergence, probability measure of fuzzy events and image thresholding. Pattern Recognition Letters 13, 857–867 (2000)
2. Bhandari, D., Pal, N.R.: Some new information measures for fuzzy sets. Information Sciences 67, 209–228 (1993)
3. Fan, J.L., Xie, W.X.: Distance measure and induced fuzzy entropy. Fuzzy Sets and systems 104, 305–314 (1999)
4. Charia, T., Ray, A.K.: Segmentation using fuzzy divergence. Pattern Recognition Letters 24, 1837–1844 (2003)
5. Tsallis, C.: Possible generalization of Boltzmann-Gibbs statistics. Journal of Statistical Physics 52, 479–487 (1988)
6. Csiszar, I.: Information type measures of differences of probability distribution and indirect observations. Studia Math. Hungarica 2, 299–318 (1967)
7. Liu, X.C.: Entropy, distance measure and similarity measure of fuzzy sets and their relations. Fuzzy Sets and Systems 52, 305–318 (1992)
8. Fan, J.L.: Fuzzy Entropy Theory. Northwest University Press, Xi'an (1999) (in Chinese)
9. Fan, J.L., Ma, Y.L., Xie, W.X.: On some properties of distance measure. Fuzzy Sets and systems 117, 355–361 (2001)
10. Furuichi, S., Yanagi, K., Kuriyama, K.: Fundamental properties of Tsallis relative entropy. Journal of Statistical Physics 45, 4868–4877 (2004)
11. CsiszarI, K.J.: Information Theory: Coding Theorems for Discrete Memoryless Systems. Academic Press, New York (1981)

Strong-II N_β-Compactness in L-Topological Spaces

Chun-yan Zhang, Yan-ping Zuo, and Shi-zhong Bai

School of Mathematics and Computational science, Wuyi University,
Guangdong 529020, China
e-mail: zhangqingwell@163.com, zuoop@126.com, bai@wyu.edu.cn

Abstract. The notion of Strong-II N_β-compactness is introduced for a complete distributive De Morgan algebra. The strong-II N_β-compactness implies strong-I N_β-compactness, hence it also implies N_β-compactness, S^*-compactness and Lowen's fuzzy compactness. But it is different from semi-compactness. When $L = [0, 1]$, strong-II N_β-compactness is equivalent to semi-compactness.

Keywords: L-topology, semi-open β_a-cover; semi-open strong β_a-cover, Strong-II N_β-compactness.

1 Introduction

In [8,9], Shi introduced the S*-compactness and N_β-compactness in L-topological spaces, where L is a complete distributive De Morgan algebra. The N_β-compactness implies S*-compactness, and S*-compactness implies Lowen's fuzzy compactness[7]. In [5], we introduced the strong-I N_β-compactness in L-topological spaces. The strong-I N_β-compactness implies N_β-compactness.

In this paper, a new compactness is introduced in L-topological spaces by means of semi-open β_a-cover and semi-open strong β_a-cover, which is called strong-II N_β-compactness, where L is a complete distributive De Morgan algebra. The strong-II N_β-compactness implies strong-I N_β-compactness. But it is different from semi-compactness. When $L = [0, 1]$, the strong-II N_β-compactness is equivalent to semi-compactness.

2 Preliminaries

Throughout this paper, $(L, \vee, \wedge, ')$ is a complete De Morgan algebra, X a nonempty set. L^X is the set of all L-fuzzy sets (or L-sets for short) on X. The smallest element and the largest element in L^X are denoted by 0 and

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 369–374.
springerlink.com

1. The set of nonunit prime elements[6,11] in L is denoted by $P(L)$. The set of nonzero co-prime elements[6,11] in L is denoted by $M(L)$. The greatest minimal family[6,11] of $a \in L$ is denoted by $\beta(a)$. An L-topological space denotes L-ts for short. Let (X, δ) be an L-ts and $A \in L^X$. Then A is called a semi-open set[1] (strongly semi-open set[2]) iff there is a $B \in \delta$ such that $B \leq A \leq B^-$ ($B \leq A \leq B^{-o}$), and A is called a semi-closed set[1] (strongly semi-closed set[2]) iff there is a $B \in \delta\prime$ such that $B^o \leq A \leq B$ ($B^{o-} \leq A \leq B$), where B^o and B^- are the interior and closure of B, respectively. $SO(L^X)$ and $SC(L^X)$ denote the family of semi-open sets and family of semi-closed sets of an L-ts (X, δ), respectively.

Definition 1. *([5]) Let (X, δ) be an L-ts, $a \in M(L)$ and $A \in L^X$. A family $\mu \subseteq L^X$ of strongly semi-open sets is called an SSO-β_a-cover of A if for any $x \in X$ with $a \notin \beta(A'(x))$, there exists a $B \in \mu$ such that $a \in \beta(B(x))$. μ is called an $SSSO - \beta_a$-cover of A if $a \in \beta(\bigwedge_{x \in X} (A'(x) \vee \bigvee_{B \in \mu} B(x)))$.*

Definition 2. *([5]) Let (X, δ) be an L-ts and $A \in L^X$. Then A is called strong-I N_β-compact if for any $a \in M(L)$, each $SSO - \beta_a$-cover of A has a finite subfamily which is an $SSSO - \beta_a$-cover of A. (X, δ) is said to be strong-I N_β-compact if 1_X is strong-I N_β-compact.*

Lemma 1. *([9]). Let L be a complete Heyting algebra, $f : X \to Y$ be a map and $f_L^{\to} : L^X \to L^Y$ is the extension of f. Then for any family $\psi \subseteq L^Y$, $\bigvee_{y \in Y} (f_L^{\to}(A)(y) \wedge \bigwedge_{B \in \psi} B(y)) = \bigvee_{x \in X} (A(x) \wedge \bigwedge_{B \in \psi} f_L^{\leftarrow}(B)(x))$.*

3 Strong-II N_β-Compactness

Definition 3. *Let (X, δ) be an L-ts, $a \in M(L)$ and $A \in L^X$. A family $\mu \subseteq SO(L^X)$ is called a semi-open β_a-cover of A if for any $x \in X$ with $a \notin \beta(A'(x))$, there exists a $B \in \mu$ such that $a \in \beta(B(x))$. μ is called a semi-open strong β_a-cover of A if $a \in \beta(\bigwedge_{x \in X} (A'(x) \vee \bigvee_{B \in \mu} B(x)))$.*

Definition 4. *Let (X, δ) be an L-ts and $A \in L^X$. Then A is called strong-II N_β-compact if for any $a \in M(L)$, each semi-open β_a-cover of A has a finite subfamily which is a semi-open strong β_a-cover of A. (X, δ) is said to be strong-II N_β-compact if 1_X is strong-II N_β-compact.*

Theorem 1. *If A is strong-II N_β-compact and B is semi-closed, then $A \wedge B$ is strong-II N_β-compact.*

Proof. Suppose that μ is a semi-open β_a-cover of $A \wedge B$. Then $\mu \cup \{B'\}$ is a semi-open β_a-cover of A. By strong-II N_β-compact of A, we know that $\mu \cup \{B'\}$ has a finite subfamily ν which is a semi-open strong β_a-cover of A. Then $\nu \setminus \{B'\}$ is a semi-open strong β_a-cover of $A \wedge B$. This shows that $A \wedge B$ is strong-II N_β-compact.

Theorem 2. *If A and B are strong-II N_β-compact in (X,δ), then $A \vee B$ is strong-II N_β-compact.*

Proof. This can be easily proved by Definition 4.

Definition 5. *Let (X,δ) and (Y,τ) be two L-ts's. A mapping $f:(X,\delta) \longrightarrow (Y,\tau)$ is called:*

(1) Irresolute [2] iff $f_L^{\leftarrow}(B)$ is semi-open in (X,δ) for each semi-open set B in (Y,τ).

(2) Semi-continuous[1] iff $f_L^{\leftarrow}(B)$ is semi-open in (X,δ) for each $B \in \tau$.

(3) Weakly irresolute iff $f_L^{\leftarrow}(B)$ is semi-open in (X,δ) for each strongly semi-open set B in (Y,τ).

Theorem 3. *If A is strong-II N_β-compact in (X,δ) and $f:(X,\delta) \longrightarrow (Y,\tau)$ is irresolute, then $f(A)$ is strong-II N_β-compact in (Y,τ).*

Proof. Let $\mu \subset SO(L^Y)$ be a semi-open β_a-cover of $f_L^{\rightarrow}(A)$. Then for any $y \in Y$, we have that

$$a \in \beta((f_L^{\rightarrow}(A))'(y) \vee \bigvee_{B\in\mu} B(y)).$$

Hence for any $x \in X$, it follows that

$$a \in \beta(A'(x) \vee \bigvee_{B\in\mu} f_L^{\leftarrow}B(x)).$$

This shows that $f_L^{\leftarrow}(\mu) = \{f_L^{\leftarrow}(B) : B \in \mu\}$ is a semi-open β_a-cover of A. By strong-II N_β-compactness of A, we know that μ has a finite subfamily ν such that $f_L^{\leftarrow}(\nu) = \{f_L^{\leftarrow}(B) : B \in \nu\}$ is a semi-open strong β_a-cover of A. By the following equation we can obtain that ν is a semi-open strong β_a-cover of $f_L^{\rightarrow}(A)$.

$$\begin{aligned} &\bigwedge_{y\in Y} ((f_L^{\rightarrow}(A))'(y) \vee \bigvee_{B\in\nu} B(y)) \\ &= \bigwedge_{y\in Y} ((\bigwedge_{x\in f^{-1}(y)} (A')(x)) \vee \bigvee_{B\in\nu} B(y)) \\ &= \bigwedge_{y\in Y} (\bigwedge_{x\in f^{-1}(y)} (A'(x) \vee \bigvee_{B\in\nu} B(f(x)))) \\ &= \bigwedge_{y\in Y} \bigwedge_{x\in f^{-1}(y)} (A'(x) \vee \bigvee_{B\in\nu} f_L^{\leftarrow}(B)(x)) \\ &= \bigwedge_{x\in X} (A'(x) \vee \bigvee_{B\in\nu} f_L^{\leftarrow}(B)(x)). \end{aligned}$$

Therefore $f_L^{\rightarrow}(A)$ is strong-II N_β-compact.

Theorem 4. *If A is strong-II N_β-compact in (X,δ) and $f:(X,\delta) \longrightarrow (Y,\tau)$ is semi-continuous, then $f(A)$ is N_β-compact in (Y,τ).*

Proof. By using definitions of N_β-compact and semi-continuous this is similar to the Theorem 3.

Theorem 5. *If A is strong-II N_β-compact in (X,δ) and $f:(X,\delta) \longrightarrow (Y,\tau)$ is weakly irresolute, then $f(A)$ is strong-I N_β-compact in (Y,τ).*

Proof. Let $\mu \subseteq L^X$ de a family of strongly semi-open sets, and μ be an SSO-β_a-cover of $f_L^{\rightarrow}(A)$. Then for any $y \in Y$, we have that $a \in \beta((f_L^{\rightarrow}(A))'(y) \vee \bigvee_{B\in\mu} B(y))$. Hence for any $x \in X$, it follows that $a \in \beta(A'(x) \vee \bigvee_{B\in\mu} f_L^{\leftarrow}B(x))$. This shows that $f_L^{\leftarrow}(\mu) = \{f_L^{\leftarrow}(B) : B \in \mu\}$ is a semi-open β_a-cover of A. By strong-II N_β-compactness of A, we know that μ has a finite subfamily ν such that $f_L^{\leftarrow}(\nu) = \{f_L^{\leftarrow}(B) : B \in \nu\}$ is a semi-open strong β_a-cover of A. Since

$$\bigwedge_{y\in Y} ((f_L^{\rightarrow}(A))'(y) \vee \bigvee_{B\in\nu} B(y)) = \bigwedge_{y\in Y} ((\bigwedge_{x\in f^{-1}(y)} (A')(x)) \vee \bigvee_{B\in\nu} B(y))$$
$$= \bigwedge_{y\in Y} (\bigwedge_{x\in f^{-1}(y)} (A'(x) \vee \bigvee_{B\in\nu} B(f(x)))) = \bigwedge_{y\in Y} \bigwedge_{x\in f^{-1}(y)} (A'(x) \vee \bigvee_{B\in\nu} f_L^{\leftarrow}(B)(x))$$
$$= \bigwedge_{x\in X} (A'(x) \vee \bigvee_{B\in\nu} f_L^{\leftarrow}(B)(x)).$$

ν is an $SSO-\beta_a$-cover of $f_L^{\rightarrow}(A)$. Therefore $f_L^{\rightarrow}(A)$ is strong-I N_β-compact.

Remark 1. The product L-ts (X, δ) of a family $\{(X_t, \delta_t)\}_{t\in T}$ of strong-II N_β-compact L-ts is not necessarily strong-II N_β-compact, whether T is a finite set or not. This can be seen from the Remark 3.8 and the Example 3.9 in [5] and the following Corollary 1.

4 Relations between Strong-II N_β-Compactness and Other Compactness

It is obvious that μ is a semi-open β_a-cover of A iff for any $x \in X$, it follows that $a \in \beta(A'(x) \vee \bigvee_{B\in\mu} B(x))$. Hence a semi-open strong β_a-cover of A is a semi-open β_a-cover of A. And a semi-open strong β_a-cover of A is an $SSO-\beta_a$-cover[5] of A, a semi-open β_a-cover of A is an $SSO - \beta_a$-cover[5] of A. When $L = [0, 1]$, μ is a semi-open β_a-cover of A iff μ' is an $a' - semi - remote$ neighborhood family[4] of A. μ is a semi-open strong β_a-cover of A iff μ' is an $(a')^- - semi - remote$ neighborhood family[4] of A. From above, we can obtain the following corollaries.

Corollary 1. *The strong-II N_β-compactness implies strong-I N_β-compactness, hence it also implies N_β-compactness, S^*-compactness and fuzzy compactness.*

Corollary 2. *When $L = [0, 1]$, strong-II N_β-compactness is equivalent to semi-compactness.*

We can easily prove the following two theorems by the Definition 4.

Theorem 6. *If (X, δ) is a regular L-ts(i.e. it satisfies $B^- = B$ for any $B \in \delta$), then an L-set $A \in L^X$ is strong-II N_β-compact iff A is strong-I N_β-compact iff A is N_β-compact.*

Theorem 7. *If for any $a, b \in L$, and $\beta(a \wedge b) = \beta(a) \cap \beta(b)$, then an L-set with a finite support is strong-II N_β-compact.*

Remark 2. (1) In general, the semi-compactness need not imply strong-II N_β-compactness. This can be seen from the Example 1.

(2) The strong-I N_β-compactness need not imply strong-II N_β-compactness.

(3) In general, if $\beta(a \wedge b) \neq \beta(a) \cap \beta(b)$, then an L-set with a finite support need not be strong-II N_β-compact. This can be seen from the Example 1.

Example 1. Let $X = \{x, y\}$ and $L = [0, 1/5] \cup \{a, b\} \cup [4/5, 1]$, where a, b are incomparable and $a \wedge b = 1/5, a \vee b = 4/5$. For each $e \in L$ with $e \neq a, b$, define $e' = 1 - e$, and $a' = b, b' = a$. Then L is a completely distributive De Morgan algebra, and

$M(L) = (0, 1/5] \cup \{a, b\} \cup (4/5, 1]$,

$\beta(a \wedge b) = \beta(1/5) = [0, 1/5) \neq [0, 1/5] = \beta(a) \cap \beta(b)$.

Take $\delta = \{0_X, A, 1_X\}$, where $A(x) = a, A(y) = b$. Then (X, δ) is an L-ts. Let $\mu = \{A\} \subset SO(L^X)$. For each $D < A$ and $D \neq 0_X$, we easily obtain that $D \notin SO(L^X)$, so μ has two subfamily $\emptyset$ and μ. Let $c = 1/5$. Obviously $c \in \beta(A(x))$ and $c \in \beta(A(y))$, this shows that μ is a semi-open β_c-cover of 1_X. But for any $\nu \subset \mu$ we have that

$c \notin \beta([\bigvee_{B \in \nu} B(x)] \wedge [\bigvee_{B \in \nu} B(y)])$.

i.e., any subfamily of μ is not a semi-open strong β_c-cover of 1_X. Therefore (X, δ) is not strong-II N_β-compact. But we have that (X, δ) is semi-compact by the Corollary 3.2 and Corollary 5.5 in [4].

Acknowledgements. This work is supported by the NSFC of China and NSF of Guangdong Province (No. 10971125, 8152902001000004).

References

1. Azad, K.K.: On fuzzy semicontinuity, fuzzy almost continuity and fuzzy weak continuity. Journal Mathematics Analysis Application 82, 14–32 (1981)
2. Bai, S.Z.: Fuzzy strongly semiopen sets and fuzzy strong semicontinuity. Fuzzy Sets and Systems 52, 345–351 (1992)
3. Bai, S.Z.: The SR-compactness in L-fuzzy topological spaces. Fuzzy Sets and Systems 87, 219–225 (1997)
4. Bai, S.Z.: The semi-precompactrness axiom. Information Sciences 165, 139–148 (2004)
5. Bai, S.Z., Fan, L., Wang, W.L.: Strong-I N_β-compactness in L-topological spaces. In: Fifth International Conference on Fuzzy Systems and Knowledge Discovery, vol. 1, pp. 436–438 (2008)
6. Liu, Y.M., Luo, M.K.: Fuzzy Topology. World Scientic, Singapore (1997)
7. Lowen, R.: Fuzzy topological spaces and fuzzy compactness. Journal Mathematics Analysis Application 56, 621–633 (1976)
8. Shi, F.G.: N_β-compactness in L-topological spaces. In: IFSA 2005, pp. 268–272. Tsinghua University Press, Beijing (2005)

9. Shi, F.G.: A new notion of fuzzy compactness in L-topological spaces. Information Sciences 173, 35–48 (2005)
10. Wang, G.J.: A new fuzzy compactness de.ned by fuzzy nets. Journal Mathematics Analysis Application 94, 1–23 (1983)
11. Wang, G.J.: Theory of L-Fuzzy Topological Space. Shaanxi Normal University Press, Xian (1988)

Some Notes on K-Harmonic Means Clustering Algorithm

Xiao-bin Zhi[1] and Jiu-lun Fan[2]

[1] School of Science, Xi' an University of Post and Telecommunications,
Xi' an 710121, China
e-mail: xbzhi@163.com
[2] School of Communication and Information Engineering,
Xi' an University of Post and Telecommunications,
Xi' an 710121, China
e-mail: jiulunf@xupt.edu.cn

Abstract. For K-harmonic means(KHM) clustering algorithm and its generalized form: KHM_P. clustering algorithm, fuzzy c-means clustering algorithm (FCM) and its generalized form: $GFCM_P$ clustering algorithms, the relations between KHM and FCM, KHM_P and $GFCM_P$ are studied. By using the reformulation of the $GFCM_P$, the facts that KHM_P is a special case of FCM_P as fuzzy parameter m is 2 and parameter p is greater than 2, and KHM is FCM as fuzzy parameter m is 2 are revealed. By using the theory of Robust Statistics, the performances of FCM_P under different parameter p is studied and the conclusions are obtained: $GFCM_p$ is sensitive to noise when parameter p is greater than 1; it is robust to noise when p is less than 1. Experimental results show the correctness of our analysis.

Keywords: Hard c-means clustering, Fuzzy c-means clustering, K-harmonic means clustering.

1 Introduction

Clustering analysis is an important branch in unsupervised pattern recognition. HCM [1] is one of the few most classical ones among the many clustering algorithms and it has gotten popular attention from scholars [2,3]. Considering HCM can not make full use of information of category, Bezdek proposed FCM clustering algorithm [4]. FCM had shown better performance over HCM in real application. However, FCM also has many drawbacks: there are almost equal numbers of data points in the clusters, almost no data points have a membership value of 1 and FCM is sensitive to noise. Therefore, many extensions to the FCM algorithm have been proposed in the literatures [5-7], Yu summarized all kinds of extensions of the FCM and proposed generalization of FCM [3,7].

Recently, As a improvement of HCM, KHM [8] and its generalized form: KHM_p [9] are proposed by Zhang capturing widely attention [10-13]. Compared with HCM, The most advantage of KHM is it enhance the robustness of algorithm

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 375–384.
springerlink.com

to initialization. KHM_P is the generalized form of KHM as the Euclidean norm is generalized as the pth power (p≥2) of the Euclidean norm in its objective function. Compared with KHM, The performance of KHM_p is further highlighted. Based on the KHM_P, Hamerly [10] presented two variants of KHM_p and compared them with HCM, FCM and gaussian mixture model based on EM (Expectation Maximization) algorithm. However, Hamerly's work is experimental in nature and lack rigorous theoretical analysis. Recently, inspired by KHM_p, Nock [11] borrowed the idea of boost in supervised classification and apply it in clustering and proposed general weighting clustering algorithm. However, Nock did not analyse the substance of KHM_p further. Other scholars [12-13] only generalized or applied KHM_P directly and could not give analysis on KHM_P theoretically.

KHM_P is really a new algorithm? FCM and KHM_p (KHM is concluded in) are improvements of HCM from different points of view, what relationship is between them? The main contribution of the paper is to answer theses questions . The remainder of this paper is organized as follows. In Section 2, FCM and KHM_P clustering algorithms are reviewed. In Section 3, the generalization of FCM: $GFCM_P$ is focused on and the reformulation of $GFCM_P$ is present. Consequently, a conclusion is arrived: KHM and KHM_P are not new clustering algorithms, that is, KHM_P is a special case of $GFCM_P$ as fuzzy parameter *m* is 2 and parameter *p* is greater than 2 and KHM is FCM as fuzzy parameter *m* is 2. Then the robust properties of $GFCM_P$ and the effects of the parameter *p* are analysed. In Section 4, some numerical examples are used to show the robustness of $GFCM_P$ when *p* is no more than 1. Finally, the conclusions are made in Section 5.

2 FCM and KHM_p

2.1 FCM

For given data set $X=\{x_1,x_2,\cdots,x_n\}$ in Euclidean space $\mathbb{R}^k$, the objective function of FCM can be written as follows

$$J_{\mathrm{FCM}}(U,V)=\sum_{i=1}^{n}\sum_{j=1}^{c}u_{ij}^{m}\left\|x_i-v_j\right\|^2 \quad , \tag{1}$$

where $\sum_{j=1}^{c}u_{ij}=1$, $1\le i\le n$, $u_{ij}\in[0,1]$, $0<\sum_{i=1}^{n}u_{ij}<n$, $1\le i\le n$, $1\le j\le c$. The necessary conditions for minimizing $J_{\mathrm{FCM}}(U,V)$ are the following update equations:

$$u_{ij}=\frac{1}{\sum_{l=1}^{c}\left(\frac{\left\|x_i-v_j\right\|^2}{\left\|x_i-v_l\right\|^2}\right)^{\frac{1}{m-1}}} \quad , \tag{2}$$

$$v_j = \frac{\sum_{i=1}^{n} u_{ij}^m \left\| x_i - v_j \right\| x_i}{\sum_{i=1}^{n} u_{ij}^m \left\| x_i - v_l \right\|} \quad , \tag{3}$$

FCM minimize the $J_{\mathrm{FCM}}(U,V)$ by alternately updating equations (2) and (3). Different from HCM, FCM adopt soft learning rule, viz. winner-takes-more, the soft learning rule mitigate the dependence of algorithm on initialization and FCM has better property of being robust to initialization.

2.2 KHM and $\mathrm{KHM_P}$

KHM and $\mathrm{KHM_P}$ are two clustering algorithms proposed by Zhang [8,9] successively, $\mathrm{KHM_P}$ is the general form of KHM, KHM is a particular case of $\mathrm{KHM_P}$ as parameter p is 2. The object function of $\mathrm{KHM_P}$ is expressed as:

$$J_{\mathrm{KHM_p}}(V) = \sum_{i=1}^{n} \frac{c}{\sum_{j=1}^{c} \frac{1}{\left\| x_i - v_j \right\|^p}} \quad , \tag{4}$$

$\mathrm{KHM_P}$ can be solved by using the fixed-point iterative method [11]. $\mathrm{KHM_P}$ is seemly a new clustering algorithm, however, it will be pointed out in Section 3, $\mathrm{KHM_P}$ is only a special case of a kind of generalized FCM and KHM is a special case of FCM.

3 The Relationship between General FCM($\mathrm{GFCM_P}$) and $\mathrm{KHM_p}$ and the Analysis on Robustness Property of $\mathrm{GFCM_P}$

3.1 General FCM: $\mathrm{GFCM_P}$

In 2005, Yu[7] presented general form of generalized FCM clustering algorithm: GFCM. The object function of GFCM is

$$J_{\mathrm{GFCM}}(U,V) = \sum_{i=1}^{n} \sum_{j=1}^{c} u_{ij}^m \rho_j(d(x_i, v_j)) - \frac{\gamma}{c} \sum_{t=1}^{c} \rho_0(d(v_j, v_t)) \quad , \tag{6}$$

where $\sum_{j=1}^{c} u_{ij} = f_i$ for $f_i \geq 0$, $\rho_j(x)$ is a continuous function of $x \in [0, +\infty)$ satisfying its derivative $\rho'_j(x) > 0$ for all $x \in [0, +\infty)$ and $\gamma \geq 0$.

Setting $f_i = 1$, $\rho_j(x) = x^p$, $\gamma = 0$, $d(x_i, v_j) = \|x_i - v_j\|$, a simple form of GFCM: GFCM$_P$ can be obtained. The object function of GFCM$_P$ is

$$J_{\mathrm{GFCM_p}}(U,V) = \sum_{i=1}^{n} \sum_{j=1}^{c} u_{ij}^m \left\| x_i - v_j \right\|^p \quad . \tag{7}$$

By Lagrange multiplier, the necessary conditions for a minimum of $J_{\mathrm{GFCM_p}}(U,V)$ can be obtained as follows:

$$u_{ij} = \frac{1}{\sum_{l=1}^{c} \left(\frac{\left\| x_i - v_j \right\|^p}{\left\| x_i - v_l \right\|^p} \right)^{\frac{1}{m-1}}} \quad , \tag{8}$$

$$v_j = \frac{\sum_{i=1}^{n} u_{ij}^m \left\| x_i - v_j \right\|^{p-2} x_i}{\sum_{i=1}^{n} u_{ij}^m \left\| x_i - v_j \right\|^{p-2}} \quad . \tag{9}$$

The iteration with update equations (8) and (9) is called the GFCM$_p$ algorithm.

3.2 The Reformulation of GFCM$_P$

For FCM, Substituting (2) into (1), a equivalent objective function with $J_{\mathrm{FCM}}(U,V)$ can be obtained as follows:

$$J_{\mathrm{RFCM}}(V) = \sum_{i=1}^{n} \left(\sum_{j=1}^{c} \left(\left\| x_i - v_j \right\|^2 \right)^{1/1-m} \right)^{1-m} \quad . \tag{10}$$

This conclusion is obtained by Hathaway as The Reformation Theorem[14]. In fact, the conclusion in Reformation Theorem can be generalized as Corollary as bellow, one can demonstrate it by similar means to theorem above, so only the conclusion is given here.

Corollary: Substituting (8) into (7), a new objective function of GFCM$_p$ can be obtained as follows

$$\min J_{\mathrm{RGFCM_p}}(V) = \sum_{i=1}^{n} \left(\sum_{j=1}^{c} \left(\left\| x_i - v_j \right\|^p \right)^{1/1-m} \right)^{1-m} \quad . \tag{11}$$

Let B is an open subset of $\Re^{ck}$, V^* *satisfying* $\left\| x_i - v^* \right\|^P > 0 \ (i = 1, 2, \cdots, n)$. *Then*

(1) (U^*, V^*) *globally minimizes* $J_{\mathrm{GFCM_p}}(U, V)$ *if and only if* V^* *globally minimizes* $J_{\mathrm{RGFCM_P}}(V)$;

(2) (U^*, V^*) *locally minimizes* $J_{\mathrm{GFCM_p}}(U, V)$ *if and only if* V^* *locally minimizes* $J_{\mathrm{RGFCM_p}}(V)$.

In (11), setting *m* to 2, a new objective function can be gotten as following form:

$$J_{\mathrm{RGFCM_p}}(V) = \sum_{i=1}^{n} \frac{1}{\sum_{j=1}^{c} \frac{1}{\left\| x_i - v_j \right\|^p}} \quad . \tag{12}$$

Comparing equation (12) with equation (4), it can be seen that equation (12) is equivalent to equation (4) except for the constant *c*. Therefore, $\mathrm{KHM_P}$ and KHM are not new clustering algorithms, $\mathrm{KHM_P}$ is in fact a special case of $\mathrm{GFCM_P}$ as fuzzy parameter *m* is 2 and parameter *p* is greater than 2, and KHM is FCM as fuzzy parameter *m* is 2. In[9], Zhang pointed out KHM is more robust to the initialization than HCM, but he did not give a interpretation theoretically. Based on the analysis above, it is known that because KHM adapt soft learning rule, so it is more robust to the initialization than HCM.

3.3 Analysis on the Robust Property of GFCMP for Different Value of Parameter p

In this Section, the property of $\mathrm{GFCM_P}$ for different value of parameter *p* is analyzed and $\mathrm{KHM_P}$ is further analysed from the $\mathrm{GFCM_P}$ point of view. For simplicity, only the situation fuzzy parameter *m* is 2 is considered here. In this case, By the equation in (9), the expression of cluster center of $\mathrm{GFCM_P}$ can be obtained as follows:

$$v_j = \frac{\sum_{i=1}^{n} u_{ij}^2 \left\| x_i - v_j \right\|^{p-2} x_i}{\sum_{i=1}^{n} u_{ij}^2 \left\| x_i - v_j \right\|^{p-2}} \quad , \tag{13}$$

where $\left\| x_i - v_j \right\|^{p-2}$ can be thought as a weighting function of x_i and it can be known

(1) When $p=2$, $\left\|x_i - v_j\right\|^{p-2} = 1$, equations in (13) will degenerate into the formula of the cluster center of FCM; (2) When $p>2$, $\lim_{\left\|x_i - v_j\right\| \to \infty} \left\|x_i - v_j\right\|^{p-2} = \infty$; (3) When $p<2$, $\lim_{\left\|x_i - v_j\right\| \to \infty} \left\|x_i - v_j\right\|^{p-2} = 0$.

According to the fact in (2), it can be known compared with FCM, KHM_P ($GFCM_P$ when the value of parameter p is more then 2) add a weighting function to data point , the weighting function is large when data point is far away from the cluster center, namely the data points locating the margin of class can obtain larger weighting value so that these data points can get more attention from the algorithm and their accuracy of clustering can be guaranteed as much as possible when algorithm is running. However, since the outlier is far away from cluster centers, the larger the value of p is, the larger the influence of the outlier is on cluster centers, then the algorithm is sensitive to outlier; the smaller the value of p is, the influence of the outlier is weaker, then the algorithm is robust to outlier. This fact can be explained rigorously from the Robust Statistics point of view. Let $[x_1, x_2, \cdots, x_n]$ be observed data set and θ is an unknown parameter to be estimated, According to the theory of Robust Statistics [15], an M-estimator of θ is generated by minimizing the form

$$\sum_{i=1}^{n} \rho(x_i, \theta) \tag{14}$$

where ρ is an arbitrary function that can measure the loss of x_i and θ. In a location estimate, the form of $\sum_{i=1}^{n} \rho(x_i - \theta)$ can be adopted and the M-estimator of θ is generated by solving the equation

$$\sum_{i=1}^{n} \varphi(x_i - \theta) = 0 \quad , \tag{15}$$

where $\varphi(x_i - \theta) = \dfrac{\partial \rho(x_i - \theta)}{\partial \theta}$. If $\rho(x-\theta)$ is taken as $\left\|x-\theta\right\|^2$, the M-estimator is the sample mean. In the M-estimator, the relative influence of individual observations toward the value of an estimate is close to the character of φ used. If φ is unbounded, the influence of outlier and noise on estimate is larger, the estimate is not accurate; If φ is bounded, the influence of outlier and noise to estimate is smaller, the estimate is accurate. If $\rho(x-\theta)$ is taken as $\left\|x-\theta\right\|^p$, the φ is $\varphi(x-\theta) = -p\left\|x-\theta\right\|^{p-2}(x-\theta)$. Then (1) If $p<1$, $\lim_{\|x\|\to\infty}\left\|\varphi(x-\theta)\right\| = \lim_{\|x\|\to\infty} p\left\|x-\theta\right\|^{p-1} = 0$, namely, φ is bounded; (2) If $p=1$, $\lim_{\|x\|\to\infty}\left\|\varphi(x-\theta)\right\| = p$, φ is also bounded; **(3)** If $p>1$, $\lim_{\|x\|\to\infty}\left\|\varphi(x-\theta)\right\| = \infty$, φ is unbounded.

The object function of $\mathrm{GFCM_P}$ can be rewritten as

$$J_{\mathrm{GFCM_p}}(U,V) = \sum_{j=1}^{c}\left[\sum_{i=1}^{n} u_{ij}^{2}\left(\left\|x_i - v_j\right\|\right)^{p}\right] \quad . \tag{16}$$

The part $\left[\sum_{i=1}^{n} u_{ij}^{2}\left(\left\|x_i - v_j\right\|\right)^{p}\right]$ in the right side of equation (16) can be thought as an M-estimator of cluster center v_j , and $[x_1, x_2, \cdots, x_N]$ is observed data set, $\rho(x-\theta) = \left\|x-\theta\right\|^{p}$ is loss function, u_{ij}^{2} is weighting function. Based on the analysis above, when $p \le 1$, $\mathrm{GFCM_P}$ would be show better performance of robust to outlier and noise; when $p > 1$, $\mathrm{GFCM_P}$ would be more sensitive to outlier and noise. Therefore, From the Robust Statistics point of view, $\mathrm{GFCM_P}$ with parameter p no more than 1 is not sensitive to outlier and noise.

4 Numerical Experiments and Analysis

In order to analyse the property of $\mathrm{GFCM_P}$, some experiments on two synthetic data sets and two real data sets are carried out in this section. In all experiment, m is set to 2. In [9], the value of p is selected as 3.5. When p is 2, $\mathrm{GFCM_P}$ is just FCM, so, three representative values of p: 2, 3.5 and 1 are selected, corresponding algorithms are denoted as FCM, $\mathrm{KHM_P}$ and $\mathrm{GFCM_1}$. we also compare $\mathrm{GFCM_P}$ with the Alternative fuzzy c-means(AFCM) clustering algorithm[5], which is proposed by Wu based on Robust Statistics to improving the robustness of FCM. The iterations were stopped as soon as the Frobenius norm in a successive pair of U matrices is less than 10^{-5}.

Synthetic Data 1

Synthetic data 1 consists of two classes strip data points shown as in Figure.1, denoted as data1. The cluster results of FCM, $\mathrm{KHM_P}$, $\mathrm{GFCM_1}$ and AFCM for data1 are shown in Fig.1 (a), Fig.1 (b), Fig.1 (c) and Fig.1 (d), respectively. FCM

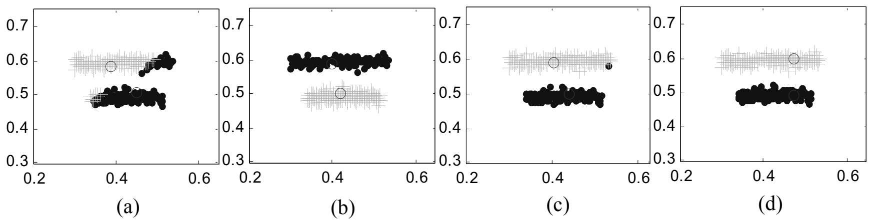


Fig. 1. Clustering results for FCM, KHM_P, $GFCM_1$ and AFCM with data1

has some misclassified data points. KHM_P and $GFCM_1$ both can classify the two clusters correctly. But we have to point out that the principles of them are not the same. Based on the analysis in section 3.3, it can be seen that KHM_P classify the data1 correctly by enhancing the separation between the clusters, however, $GFCM_1$ weaken the influence of the data points located at the margin and can also cluster the data1 correctly. AFCM also show robustness to data1 and cluster data1 correctly, but the cluster centers have the deviation on some degree.

Synthetic Data 2

Synthetic2 consists of data1 with 50 Gaussian noise data points, denoted as data2. The cluster results of FCM, KHM_P, $GFCM_1$ and AFCM for data3 are shown in Fig.2 (a), Fig.2 (b), Fig.2 (c) and Fig.2 (d) respectively. FCM and KHM_P is sensitive to noise and can not cluster the data2 correctly, however, $RGFCM_1$ and AFCM show robustness to noise and both can cluster the data2 correctly. The cluster centers of $GFCM_1$ are relatively accurate and the cluster centers of AFCM have the deviation on some degree.

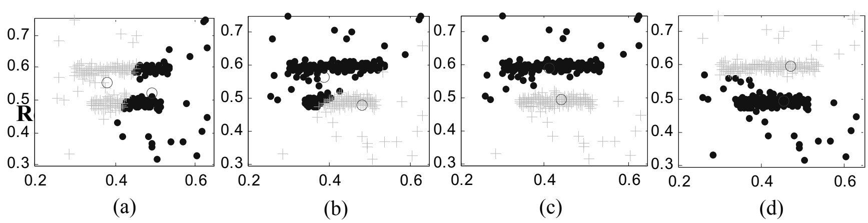


Fig. 2. Clustering results of FCM, KHM_P, $GFCM_1$ and AFCM for data2

Real world data

Iris and Wine both are retrieved from the UCI repository of machine learning databases[16] and their properties are listed in table1. The total error counts of the four clustering algorithms using the two datasets are shown in table 2.

Table 1. Dataset Descriptions

	Size of dataset	Number of dimensions	Number of clusters
Iris	150	4	3
Wine	178	13	3

Table 2. Clustering results for FCM, KHM_P, $GFCM_1$ and AFCM with data1

	FCM	KHM_P	$GFCM_1$	AFCM
Iris	16	16	11	11
Wine	56	72	50	54

Based on the table 2, it can be known that for the Iris data, the total error counts for FCM and KHM both are 16, however, the total error counts for $GFCM_1$ and AFCM is only 11. Therefore, for the Iris data, FCM_P and AFCM show superior performance over FCM and KHM. For the Wine data, The total error counts for FCM, KHM_P, $GFCM_1$ and AFCM are respectively 56, 72, 50 and 54. $GFCM_1$ shows the best clustering performance in these four clustering algorithms. To sum up above analysis, due to the complication of real data sets, robust clustering algorithm can obtain good performance. According to the experiment results above, it can be known that $GFCM_1$ has superior performance over other three clustering algorithms.

5 Conclusion

KHM clustering algorithm and KHM_P clustering algorithm obtain widely attention, but it have been shown that KHM and KHM_P are not new clustering algorithms. By using the reformulation of the $GFCM_P$, it is pointed out that KHM_P is a special case of $GFCM_P$ as fuzzy parameter m is 2 and parameter p is greater than 2, and KHM is FCM when fuzzy parameter m is 2. Additionally, the performances of $GFCM_P$ with different parameter p are analysized and the conclusion has been obtained that as parameter p is greater than 1, $GFCM_p$ would be sensitive to outlier and noise; when p is no greater than 1, $GFCM_p$ would be robust to outlier and noise. The performance of $GFCM_p$ is depand on the choice of parameter p, how to adaptively choose parameter p is future work to be do.

References

1. McQueen, J.: Some methods for classification and analysis of multivariate observations. In: Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, pp. 281–297 (1967)

2. Selim, S.Z., Ismail, M.A.: K-Means-Type Algorithms: A Generalized Convergence Theorem and Characterization of Local Optimality. IEEE Trans. Pattern Analysis and Machine Intelligence 6(1), 81–87 (1984)
3. Jian, Y.: General C-Means Clustering Model. IEEE Trans. Pattern Analysis and Machine Intelligence 27(5), 657–667 (2005)
4. Bezdek, J.C.: Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum, New York (1981)
5. Wu, K.L., Yang, M.S.: Alternative c-means clustering algorithms. Pattern Recognition 35(10), 2267–2278 (2002)
6. Fan, J.L., Zhen, W.Z., Xie, W.X.: Suppressed fuzzy c-means clustering algorithm. Pattern Recognition Letters 24(9-10), 1607–1612 (2003)
7. Yu, J.: Optimality Test for Generalized FCM and Its Application to Parameter Selection. IEEE Trans. Fuzzy Systems 13(1), 164–176 (2005)
8. Zhang, B., Hsu, M., Dayal, U.: K-harmonic means—a data clustering algorithm. Technical Report HPL-124, Hewlett-Packard Labs (1999)
9. Zhang, B.: Generalized K-harmonic means—boosting in unsupervised learning. Technical Report HPL-2000-137, Hewlett-Packard Labs (2000)
10. Hamerly, G., Elkan, C.: Alternatives to k-means algorithm that find better clusterings. In: Proceedings of the eleventh international conference on Information and knowledge management, McLean, Virginia, USA, pp. 600–607 (2002)
11. Nock, R., Nielsen, F.: On Weighting Clustering. IEEE Trans. Pattern Analysis and Machine Intelligence 28(8), 1223–1235 (2006)
12. Li, Q., Mitianoudis, N., Stathaki, T.M.: Spatial kernel K-harmonic means clustering for multi-spectral image segmentation. IET, Image Processing 1(2), 156–167 (2007)
13. Güngör, Z., Ünler, A.: K-harmonic means data clustering with simulated annealing heuristic. Applied Mathematics and Computation 184, 199–209 (2007)
14. Hathaway, R.J., Bezdek, J.C.: Optimization of clustering criteria by reformulation. IEEE Trans. Fuzzy Systems 3(2), 241–246 (1995)
15. Huber, P.J.: Robust Statistics, New York (1981)
16. Blake, C.L., Merz, C.J.: UCI repository of machine learning databases (1998), `http://archive.ics.uci.edu/ml/`

On Fuzzy Prime Filters of Lattice H Implication Algebras

Yi Liu[1,2], Ling Guo[3], Xiao-yan Qin[2,4], and Yang Xu[2]

[1] Key Laboratory of Numerical Simulation in the Sichuan Provincial College, Neijiang Normal University, Neijiang 641000, Sichuan, China

[2] Intelligent Control Development Center, Southwest Jiaotong University, Chengdu 610031 Sichuan, P.R. China
e-mail: liuyiyl@126.com

[3] School of Mathematics, Southwest Jiaotong University, Chengdu 610031 Sichuan, P.R. China
e-mail: guo1987ling@163.com

[4] College of Mathematics and Computer Science, Shanxi Normal University, Linfen 041004, Shanxi, China

Abstract. In this paper,firstly, we introduce the fuzzy annihilator A^* of a fuzzy set A and its properties are obtained. Secondly, the fuzzy prime filter of lattice H implication algebras are studied by using the fuzzy annihilator. Finally, we obtain that $FPF(L_H)$, the set of all fuzzy prime filter of a lattice H implication algebra, is a bounded lattice with an order-reversing involution *.

Keywords: Lattice H implication algebra, fuzzy filter, fuzzy prime filter.

1 Introduction

Many-valued logic, a great extension and development of classical logic, has always been a crucial direction in non-classical logic. In order to research the many-valued logical system whose propositional value is given in a general lattice, in 1993, Xu firstly established the lattice implication algebra by combining lattice and implication algebra, and investigated many useful structures [1],[2],[3],[4]. Xu and Qin proposed the lattice H implication algebras (**LHIA** for short) and investigated its properties[5]. Lattice implication algebra provided the foundation to establish the corresponding logical system from the algebraic viewpoint. For the general development of lattice implication algebras, the filter theory plays an important role. Meanwhile, filter plays an important role in automated reasoning and approximated reasoning based on lattice implication algebra, too, for example, J. Ma, et al. proposed filter-based resolution principle[8]. Xu and Qin [6]introduced the notions of filter and implicative filter in a lattice implication algebra, and investigated

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 385–393.
springerlink.com

their properties[21],[22]. Y.B.Jun and other scholars studied several filters in lattice implication algebras [14],[15],[16] [19],[20],[21],[22]. Xu and Qin investigated prime filters of lattice H implication algebras[23].

The concept of fuzzy set was introduced by Zadeh (1965). Since then this idea has been applied to other algebraic structures such as groups, semigroups, rings, modules, vector spaces and topologies. Xu and Qin[2] applied the concept of fuzzy set to lattice implication algebras, they applied the concept of fuzzy sets to lattice implication algebras and proposed the notions of fuzzy filters and fuzzy implicative filters[3]. Later on, some scholars introduced related fuzzy filter such as fuzzy (positive) implication filter, fuzzy fantastic filter and investigated some properties[9],[10],[11],[12],[17],[18],[22]. This logical algebra has been extensively investigated by several researchers, and many elegant results are obtained, collected in the monograph[4].

In this paper, as an extension of fuzzy filters theory in lattice implication algebras, we further study the fuzzy filters of lattice H implication algebras. In Section 2, we list some preliminaries, which are useful to development this topic in other sections. In section 3, we first introduce fuzzy annihilator of fuzzy set and investigate its properties; the properties of fuzzy prime filter on lattice H lattice implication algebras are investigated and obtain that $FPF(L_H)$, the set of all fuzzy prime filter of a lattice H implication algebra, is a bounded lattice with an order-reversing involution *.

In this paper denote $\mathscr{L}$ as lattice (resp. lattice H) implication algebra $(L, \vee, \wedge, ', \rightarrow, O, I)$.

2 Preliminaries

Definition 1. *[1] Let $(L, \vee, \wedge, O, I)$ be a bounded lattice with an order-reversing involution $'$. The greatest element I and the smallest element O, and*

$$\rightarrow: L \times L \longrightarrow L$$

be a mapping. $\mathscr{L} = (L, \vee, \wedge, ', \rightarrow, O, I)$ is called a lattice implication algebra if the following conditions hold for any $x, y, z \in L$:

(I_1) $x \rightarrow (y \rightarrow z) = y \rightarrow (x \rightarrow z)$;
(I_2) $x \rightarrow x = I$;
(I_3) $x \rightarrow y = y' \rightarrow x'$;
(I_4) $x \rightarrow y = y \rightarrow x = I$ implies $x = y$;
(I_5) $(x \rightarrow y) \rightarrow y = (y \rightarrow x) \rightarrow x$;
(l_1) $(x \vee y) \rightarrow z = (x \rightarrow z) \wedge (y \rightarrow z)$;
(l_2) $(x \wedge y) \rightarrow z = (x \rightarrow z) \vee (y \rightarrow z)$.

Theorem 1. *[4] Let $\mathscr{L}$ be a lattice implication algebra. Then for any $x, y, z \in L$, the following conclusions hold:*

(1) if $I \rightarrow x = I$, then $x = I$;
(2) $I \rightarrow x = x$ and $x \rightarrow O = x'$;
(3) $O \rightarrow x = I$ and $x \rightarrow I = I$;

(4) $(x \to y) \to ((y \to z) \to (x \to z)) = I$;
(5) $(x \to y) \vee (y \to x) = I$;
(6) *if* $x \le y$, *then* $x \to z \ge y \to z$ *and* $z \to x \le z \to y$;
(7) $x \le y$ *if and only if* $x \to y = I$;
(8) $(z \to x) \to (z \to y) = (x \wedge z) \to y = (x \to z) \to (x \to y)$;
(9)$x \to (y \vee z) = (y \to z) \to (x \to z)$;
(10) $x \vee y = (x \to y) \to y$;
(11) $x \vee y = I$ *if and only if* $x \to y = y$.

Definition 2. *[5] A lattice implication algebra $\mathscr{L}$ is called a lattice H implication algebra, if for any $x, y, z \in L$*

$$(x \vee y) \vee ((x \wedge y) \to z) = I.$$

Theorem 2. *[4] Let $\mathscr{L}$ be a lattice implication algebra. Then following statements are equivalent:*
(1) $\mathscr{L}$ *be a lattice H implication algebra;*
(2) For any $x, y \in L$, $x \to (x \to y) = x \to y$;
(3) For any $x, y, z \in L$, $x \to (y \to z) = (x \to y) \to (x \to z)$;
(4) For any $x, y, z \in L$, $x \to (y \to z) = (x \wedge y) \to z$;
(5) For any $x, y \in L$, $(x \to y) \to x = x$;
(6) For any $x \in L$, $x \vee x' = I$.

In lattice implication algebras, Define the binary operation $\otimes$ as follows: for any $x, y \in L$, $x \otimes y = (x \to y')'$. Some properties of operation $\otimes$ are in reference[4].

A fuzzy set in $\mathscr{L}$ is a function$A : L \to [0, 1]$. For a fuzzy set A in $\mathscr{L}$ and $t \in [0, 1]$, the set $A_t = \{x \in L | A(x) \ge t\}$ is called the level subset of A. For any fuzzy sets A and B in $\mathscr{L}$, we define $A \subseteq B$ if and only if $A(x) \le B(x)$ for any $x \in L$; $(A \cap B)(x) = min\{A(x), B(x)\}$ for any $x \in L$.

Definition 3. *[2] Let A be a fuzzy set of a lattice implication algebra $\mathscr{L}$. Then A is called a fuzzy filter of $\mathscr{L}$ if, for any $x, y \in L$*
(1) $A(I) \ge A(x)$;
(2) $A(y) \ge min\{A(x \to y), A(x)\}$.

Theorem 3. *[22] Let A be a fuzzy subset of L. Then A is a fuzzy filter if and only if , for any $x, y, z \in L$, $x \to (y \to z) = I$ implies $min\{A(x), A(y)\} \le A(z)$.*

Definition 4. *[24] A fuzzy filter A of $\mathscr{L}$ is said to be fuzzy prime if it is non-constant and $A(x \vee y) = A(x) \vee A(y)$ for any $x, y \in L$.*

3 Fuzzy Prime Filters of LHIA

In reference[4], put $F_a = \{x \in L | x \vee a = I, a \in L\}$, then F_a is a filter of $\mathscr{L}$. Let A be a fuzzy set of $\mathscr{L}$. Define a new fuzzy set as follow:

$$A^*(x) = 1 - inf\{A(y)|y \in F_x\} \tag{1}$$

for any $x \in L$. A^* is called **fuzzy annihilator** of A.

Lemma 1. *Let A and B be two fuzzy sets of a lattice implication algebra $\mathscr{L}$. Then the following statements are hold:*
(1) $A^*(I) \geq A^*(x)$ for any $x \in L$;
(2) If $x \leq y$, then $A^*(x) \leq A^*(y)$;
(3) If $A \subseteq B$, then $B^* \subseteq A^*$.

Proof. (1) For any $x \in L$, we have $x \vee I = I$.

$$\begin{aligned} A^*(I) &= 1 - inf\{A(y)|y \in F_I\} = 1 - inf\{A(y)|y \in L\} \\ &\geq 1 - inf\{A(y)|y \in F_x\} = A^*(x). \end{aligned}$$

(2) Let $z \in F_x$. Then $z \vee x = I$. Since $x \leq y$, so $y \vee z = I$, hence $z \in F_y$. It follows that

$$\{z \in L|z \in F_x\} \subseteq \{z \in L|z \in F_y\}.$$

Therefore,

$$A^*(x) = 1 - inf\{A(z) \in L|z \in F_x\} \leq 1 - inf\{A(z) \in L|z \in F_y\} = A^*(y).$$

(3) Since $A \subseteq B$, then $inf\{A(z)|z \in F_x\} \leq inf\{B(z)|z \in F_x\}$. I follows that

$$A^*(x) = 1 - inf\{A(z)|z \in F_x\} \geq 1 - inf\{B(z)|z \in F_x\} = B^*(x).$$

Hence, $B^* \subseteq A^*$.

Theorem 4. *Let $\mathscr{L}$ be a lattice H implication algebra and A be a fuzzy filter of $\mathscr{L}$. Then $(A^*)^* = A$.*

Proof. Let A be a fuzzy filter of $\mathscr{L}$. For any $x \in L$, we have:

$$\begin{aligned} (A^*)^*(x) &= 1 - inf\{A^*(y)|y \in F_x\} \\ &= 1 - inf\{1 - inf\{A(z)|z \in F_y\}|y \in F_x\} \\ &= sup_{y \in F_x}\{1 - (1 - inf\{A(z)|z \in F_y\})\} \\ &= sup_{y \in F_x}\{inf\{A(z)|z \in F_y\}\} \\ &\geq inf\{A(z)|z \in F_{x'}\}. \end{aligned}$$

Since $z \in F_{x'}$, then $z \vee x^{'} = I$, that is, $(z \to x^{'}) \to x^{'} = I$. It follows that

$$\begin{aligned} (z \to x^{'}) \to x^{'} &= (z \to x^{'}) \to (x \to O) \\ &= x \to (z^{'} \to O) = x \to z = I. \end{aligned}$$

That is, $x \leq z$. Since A is a fuzzy filter of $\mathscr{L}$, it follows that $A(x) \leq A(z)$ for any $z \in F_{x'}$. Therefore

$$(A^*)^*(x) \geq inf\{A(z)|z \in F_{x'}\} = A(x).$$

That is, $(A^*)^* \supseteq A$.

On the other hand, for any $x \in L$,

$$\begin{aligned}(A^*)^*(x) &= 1 - inf\{A^*(y)|y \in F_x\} \\ &= 1 - inf\{1 - inf\{A(z)|z \in F_y\}|y \in F_x\} \\ &\leq 1 - inf\{1 - A(x)|y \in F_x\} = A(x),\end{aligned}$$

that is, $(A^*)^* \subseteq A$. Therefore $A^*)^* = A$.

Lemma 2. *[24] Let A be a fuzzy filter of $\mathscr{L}$. Then A is a constant fuzzy set if and only if $A(I) = A(O)$.*

Lemma 3. *[24] Let A be a non-constant fuzzy filter of $\mathscr{L}$. Then the following are equivalent:*

(1) A is a fuzzy prime filter of $\mathscr{L}$,
(2) for all $x, y \in L$, if $A(x \vee y) = A(I)$, then $A(x) = A(I)$ or $A(y) = A(I)$,
(3) for all $x, y \in L$, $A(x \to y) = A(I)$ or $A(y \to x) = A(I)$.

Lemma 4. *Let A be a fuzzy prime filter of a lattice implication algebra $\mathscr{L}$. Then $L_A := \{x \in L|A(x) = A(I)\}$ is a prime filter of $\mathscr{L}$.*

Proof. Suppose that A is a fuzzy prime filter of $\mathscr{L}$. By Corollary 3.6 in[4], we have L_A is a filter.

Since A is non-constant, L_A is proper. Let $x \vee y \in L_A$ for any $x, y \in L$. Then $A(I) = A(x \vee y) = A(x) \vee A(y)$. Hence $A(x) = A(I)$ or $A(y) = A(I)$. This means that $x \in L_A$ or $y \in L_A$. Therefore, L_A is prime.

Conversely, assume that L_A is a prime filter of $\mathscr{L}$. Since L_A is proper, A is non-constant. As $(x \to y) \vee (y \to x) = I$ for any $x, y \in L$. Then $x \to y \in L_A$ or $y \to x \in L_A$. That is, $A(x \to y) = A(I)$ or $A(y \to x) = A(I)$. So A is a fuzzy prime filter of $\mathscr{L}$.

Lemma 5. *[4] Let $\mathscr{L}$ be a lattice implication algebra. Then $[x) \cap [y) = [x \vee y)$ for any $x, y \in L$.*

Theorem 5. *Let A be a fuzzy prime filter of a lattice H implication algebra $\mathscr{L}$. Then*

$$A^*(x) = \begin{cases} 1 - A(O), & if \quad x \in L_A, \\ 1 - A(I), & if \quad otherwise. \end{cases} \tag{2}$$

Proof. Suppose that A is a fuzzy prime filter. If $x \notin L_A$. Since $y \vee x = I$ for any $y \in F_x$. Thus $A(I) = A(x \vee y)$. It follows that $A(x) = A(I)$ or $A(y) = A(I)$ by Lemma 3. Since $x \notin L_A$, so $A(x) \neq A(I)$, hence $A(y) = A(I)$. Therefore, $A^*(x) = 1 - inf\{A(y)|y \in F_x\} = 1 - A(I)$.

If $x \in L_A$, and A is non-constant, then $A(x) = A(I) \neq A(O)$. By the hypothesis, A is a fuzzy filter of $\mathscr{L}$, so $A(O) \geq min\{A(x), A(x \to O)\}$ and A is order-preserving, it follows that $A(O) = min\{A(x), A(x \to O)\} = A(x^{'})$.

On the other hand, $\mathscr{L}$ is a lattice H implication algebra, then $x \vee x^{'} = I$ for any $x \in L$. That is, $x^{'} \in F_x$. Since $A(O) = A(x^{'}) \leq A(y)$ for any $y \in L$. Therefore, $A^*(x) = 1 - inf\{A(y)|y \in F_x\} = 1 - A(x^{'}) = 1 - A(O)$.

Lemma 6. *[4] Let A be a fuzzy set of $\mathscr{L}$ and $A \neq \emptyset$. Then A is a fuzzy filter of $\mathscr{L}$ if and only if for any $t \in [0,1]$, A_t is a filter of $\mathscr{L}$ when $A_t \neq \emptyset$.*

Lemma 7. *[4] Let $\mathscr{L}$ be a lattice implication algebra, J is a proper filter of $\mathscr{L}$. Then the following statements are equivalent:*

(1) J is irreducible;

(2) $[a) \cap [b) \subseteq J$ implies $a \in J$ or $b \in J$ for any $a, b \in L$;

(3) J is prime.

Theorem 6. *Let A be a fuzzy prime filter of a lattice H implication algebra $\mathscr{L}$. Then so does A^*.*

Proof. Let A be a fuzzy prime filter of a lattice H implication algebra. Now we need to prove that $A_t^* \neq \emptyset$ is a filter of $\mathscr{L}$ for any $t \in [0,1]$.

Since $A^*(I) \geq A^*(x)$ for any $x \in L$, then, for any $t \in [0,1]$ and $x \in A_t^*$, we have $A^*(I) \geq t$. That is $I \in A_t^*$. Let $x, x \to y \in A_t^*$, then $A^*(x) \geq t$ and $A^*(x \to y) \geq t$. That is,

$$A^*(x) = 1 - inf\{A(z)|z \in F_x\} \geq t,$$
$$A^*(x) = 1 - inf\{A(z)|z \in F_{x \to y}\} \geq t,$$

then,

$$inf\{A(z)|z \in F_x\} \leq 1 - t,$$
$$inf\{A(z)|z \in F_{x \to y}\} \leq 1 - t.$$

Therefore, for any $\varepsilon > 0$, there exists $z_1 \in F_x$ and $z_2 \in F_{x \to y}$ such that $A(z_1) \leq 1 - t + \varepsilon$ and $A(z_2) \leq 1 - t + \varepsilon$.

Since $z_1 \in F_x$ and $z_2 \in F_{x \to y}$, so $x \to z_1 = z_1$ and $(x \to y) \to z_2 = z_2$. It follows that

$$\begin{aligned}(x \to y) \to (z_1 \vee z_2) &= (x \to y) \to ((z_1 \to z_2) \to z_2)\\ &= (z_1 \to z_2) \to ((x \to y) \to z_2)\\ &= (z_1 \to z_2) \to z_2 = z_1 \vee z_2,\end{aligned}$$

that is, $z_1 \vee z_2 \in F_{x \to y}$. Since $x \to z_1 = z_1$, so

$$\begin{aligned}x \to (z_1 \vee z_2) &= x \to ((z_2 \to z_1) \to z_1)\\ &= (z_2 \to z_1) \to (x \to z_1)\\ &= (z_2 \to z_1) \to z_1 = z_1 \vee z_2,\end{aligned}$$

that is, $z_1 \vee z_2 \in F_x$. And so

$$\begin{aligned} z_1 \vee z_2 &= (x \to y) \to (z_1 \vee z_2) \\ &= (x \to y) \to (x \to (z_1 \vee z_2)) \\ &= x \to (y \to (z_1 \vee z_2)) \\ &= y \to (x \to (z_1 \vee z_2)) \\ &= y \to (z_1 \vee z_2). \end{aligned}$$

That is, $z_1 \vee z_2 \in F_y$.

Since A is a fuzzy prime filter of $\mathscr{L}$, we have that $A(z_1 \vee z_2) = A(z_1) \vee A(z_2) \leq 1 - t + \varepsilon$. Hence, $inf\{A(z)|z \in F_y\} \leq A(z_1 \vee z_2) \leq 1 - t + \varepsilon$. It follows that $A^*(y) = 1 - inf\{A(z)|z \in F_y\} \geq t - \varepsilon$. By the arbitrariness of ε, we have that $A^*(y) \geq t$, that is, $y \in A^*_t$. Therefore, by Lemma 6, A^*_t is a filter of $\mathscr{L}$. It follows that A^* is a fuzzy filter of $\mathscr{L}$.

Now, we need to prove A^* is prime, that is, to prove $A^*(x \vee y) = A^*(x) \vee A^*(y)$ for any $x, y \in L$.

Case (a): If $x \vee y \in L_A$, then $A^*(x \vee y) = 1 - A(O)$ by Theorem 5. Since L_A is a prime filter, so $[x \vee y) \subseteq L_A$. By Lemma 2, it follows that $[x \vee y) = [x) \cap [y) \subseteq L_A$. For L_A is prime, so $x \in L_A$ or $y \in L_A$. Therefore, $A^*(x) = 1 - A(O) = A^*(x \vee y)$ or $A^*(x) = 1 - A(O) = A^*(x \vee y)$. And so $A^*(x \vee y) = A^*(x) \vee A^*(y)$.

Case (b): If $x \vee y \notin L_A$, then $x \notin L_A$ and $y \notin L_A$ for L_A is a prime filter of $\mathscr{L}$, it follows that $A^*(x) = 1 - A(I) = A^*(y)$. Therefore $A^*(x \vee y) = A^*(x) \vee A^*(y)$.

Sum up above, A^* is a fuzzy prime filter of $\mathscr{L}$.

Lemma 8. *Let A and B are non-constant fuzzy filter of a lattice H implication algebra $\mathscr{L}$. Then A^*, $A \cap B$ and $A \cup B$ are non-constant.*

Proof. From Theorem 4, we have that $(A^*)^* = A$. If A is constant, then $A(x) = c$(c is constant) for any $x \in L$. Since

$$\begin{aligned} A(x) &= (A^*)^*(x) \\ &= 1 - inf\{A^*(y)|y \in F_x\} \\ &= 1 - inf\{c|y \in F_x\} = 1 - c, \end{aligned}$$

that is, A is constant, contradiction. Therefore A^* is a constant.

Since A and B are non-constant, then there exist $x, y \in L$ such that $A(x) < A(I)$ and $B(y) < A(I)$. Without loss of generality, we can assume that $B(y) \leq A(x)$. It follows that

$$\begin{aligned} (A \cap B)(I) &= min\{A(I), B(I)\} \\ &> min\{A(x), B(y)\} \\ &= B(y) \geq min\{A(y), B(y)\} = (A \cap B)(y). \end{aligned}$$

That is, $A \cap B$ is non-constant.

Since A and B are non-constant fuzzy filter of a lattice H implication algebra $\mathscr{L}$, then $(A^*)^* = A$ and $(B^*)^* = B$. As $A \cap B \subseteq A, B$, by Lemma 1, we have $(A \cap B)^* \supseteq A^*$ and$(A \cap B)^* \supseteq B^*$, then $(A \cap B)^* \supseteq A^* \cup B^*$.

On the other hand, $A^*, B^* \subseteq A^* \cup B^*$, so $(A^* \cup B^*)^* \subseteq (A^*)^*, (B^*)^*$, then $(A^* \cup B^*)^* \subseteq (A^*)^* \cap (B^*)^* = A \cap B$. And so $(A \cap B)^* \subseteq A^* \cup B^*$. It follows that $(A \cap B)^* = A^* \cup B^*$. Therefore, $A \cup B = (A^*)^* \cup (B^*)^* = (A^* \cap B^*)^*$. It follows that $A \cup B$ is non-constant.

Theorem 7. *Let A and B are fuzzy prime filter of a lattice H implication algebra $\mathscr{L}$. Then $A \cap B$ and $A \cup B$ are also fuzzy prime filter of $\mathscr{L}$.*

Proof. We only to prove that $A \cup B$ is a fuzzy prime filter of $\mathscr{L}$. From the proof of Lemma 8, we have that $A \cup B = (A^* \cap B^*)^*$. Since A and B are fuzzy prime filters, so A^*, B^* are fuzzy prime filters by Theorem 6. Then $A^* \cap B^*$ is fuzzy prime. By Theorem 6 again, we have $(A^* \cap B^*)^* = A \cup B$ is fuzzy prime.

Denote $FPF(L_H)$ by the set of all fuzzy prime filter of a lattice H implication algebra. From Lemma 1, Theorem 4, and Theorem 7, we have:

Theorem 8. $(FPF(L_H); \cup, \cap,)$ *is a bounded lattice with an order-reversing involution* *.

4 Conclusion

In order to research the many-valued logical system whose propositional value is given in a lattice, Xu and Qin initiated the concept of lattice H implication algebras. Hence for development of this many-valued logical system, it is needed to make clear the structure of lattice implication algebras. It is well known that to investigate the structure of an algebraic system, the filters with special properties play an important role. In this paper, we first introduced fuzzy annihilator of fuzzy set and investigated its properties; the properties of fuzzy prime filter on lattice H lattice implication algebras are investigated and obtain that $FPF(L_H)$, the set of all fuzzy prime filter of a lattice H implication algebra, is a bounded lattice with an order-reversing involution *.

Acknowledgements. This work is supported by National Natural Science Foundation of P.R.China (Grant no. 60875034).

References

1. Xu, Y.: Lattice implication algebras. J. Southwest Jiaotong Univ. 1, 20–27 (1993)
2. Xu, Y., Qin, K.Y.: Fuzzy lattice implication algebras. J. Southwest Jiaotong Univ. 2, 121–127 (1995)

3. Xu, Y., Qin, K.Y.: On filters of lattice implication algebras. J. Fuzzy Math. 2, 251–260 (1993)
4. Xu, Y., Ruan, D., Qin, K., Liu, J.: Lattice-Valued Logic-An Alternative Approach to Treat Fuzziness and Incomparability. Springer, New York (2003)
5. Xu, Y., Qin, K.Y.: Lattice H implication algebras and lattice implication algebra classes. J. Hebei Mining and Civil Engineering Institute 3, 139–143 (1992)
6. Xu, Y., Qin, K.Y.: On filters of lattice implication algebras. J. Fuzzy Math. 1(2), 251–260 (1993)
7. Xu, Y., Qin, K.Y.: Fuzzy lattice implication algebras. J. Southwest Jiaotong Univ. 30(2), 121–127 (1995)
8. Ma, J., Li, W.J., Ruan, D., Xu, Y.: Filter-based resolution principle for lattice -valued propsitional logic LP(x). Infor. Sci. 177, 1046–1062 (2007)
9. Jun, Y.B., Xu, Y., Ma, J.: Redefined fuzzy implication filters. Infor. Sci. 177, 1422–1429 (2007)
10. Zhan, J.M., Jun, Y.B.: Notes on redefined fuzzy implication filters of lattice implication algebras. Infor. Sci. 179, 3182–3186 (2009)
11. Jun, Y.B.: Fuzzy positive implicative and fuzzy associative filters of lattice implication algebras. Fuzzy Sets Syst. 121, 353–357 (2001)
12. Rom, E.H., Kim, S.Y., Xu, Y., et al.: Some operations on lattice implication algebras. Int. J. Math. Sci. 1, 45–52 (2001)
13. Jun, Y.B.: Implicative filters of lattice implication algebras. Bull. Korean Math. Soc. 34(2), 193–198 (1997)
14. Jun, Y.B.: Fantastic filters of lattice implication algebras. Internat. J. Math. and Math. Sci. 24(4), 277–281 (2000)
15. Jun, Y.B.: On n-fold implicative filters of lattice implication algebras. Int. J. Math. and Math. Sci. 26(11), 695–699 (2001)
16. Jun, Y.B.: Fuzzy positive implicative and fuzzy associative filters of lattice implication algebras. Fuzzy Sets Syst. 121, 353–357 (2001)
17. Jun, Y.B., Song, S.Z.: On fuzzy implicative filters of lattice implication algebras. J. Fuzzy Math. 10(4), 893–900 (2002)
18. Jun, Y.B., Xu, Y., Qin, K.Y.: Positive implicative and associative filters of lattice implication algebras. Bull. Korean Math. Soc. 35(1), 53–61 (1998)
19. Kim, S.Y., Roh, E.H., Jun, Y.B.: On ultra filters of lattice implication algebras. Scientiae Mathematicae 2(2), 201–204 (1999)
20. Liu, J., Xu, Y.: On certain filters in lattice implication algebras. Chinese Quarterly J. Math. 11(4), 106–111 (1996)
21. Liu, J., Xu, Y.: Filters and structure of lattice implication algebras. Chinese Science Bulletin 42(18), 1517–1520 (1997)
22. Xu, W.T., Xu, Y.: Structure of fuzzy filters in lattice implication algebras. Chinese Quar. J. Math. 6, 512–518 (2008)
23. Xu, Y., Qin, K.Y., Roh, E.H.: On prime filters of lattice H implication algebras. Int. J. Fuzzy Math. 9(3), 583–588 (2001)
24. Liu, Y., Qin, K.Y., Xu, Y.: Fuzzy prime filter of lattice implication algebras. Fuzzy information and Engineering (2010) (submitted)

Numerical Characteristics of Intuitionistic Fuzzy Sets

Zhi-ming Bai[1], Dao-wu Pei[1], and Min Mao[2]

[1] College of Science, Zhejiang Sci-Tech University,
Hangzhou 310018, China
e-mail: 395855084bzm@163.com, peidw@163.com
[2] College of Information and Electronics,
Zhejiang Sci-Tech University, Hangzhou 310018, China
e-mail: 849232058@qq.com

Abstract. In this paper, we discuss four numerical characteristics of intuitionistic fuzzy sets. These numerical characteristics include distance, similarity measure, inclusion measure and entropy of intuitionistic fuzzy sets. By constructing the corresponding formulas, we show that any one numerical characteristic of four can be expressed by any other numerical characteristic. Thus some direct connections between these numerical characteristics are set up.

Keywords: Intuitionistic fuzzy set; Distance; Similarity measure; Inclusion measure; Entropy.

1 Introduction

Since fuzzy set theory is proposed by Zadeh in 1965, some theories and methods have be proposed for processing uncertainty and imprecise. In them two theories are notable: intuitionistic fuzzy set (IFS) theory proposed by Atanasov [1] and interval-valued fuzzy set theory. Deschrijver and Kerre [5] constructed an isomorphic mapping between IVFSs and IFSs in 2003. Although there are compact interconnect between them, IVFSs and IFSs represent different information and have different semantics: IVFSs focus on the uncertainty of membership function, but IFSs emphasize the relationship between membership function and non-membership function.

Four numerical characteristics including distance, similarity measure, inclusion measure and entropy are important research objects in fuzzy set theory. Wang [8] proposed the concept of the similarity measure. Moreover it has been used to many areas such as cluster analysis, image processing, approximate reasoning, fuzzy control, etc. Inclusion measure of fuzzy sets are used to describe the degree of a fuzzy set be included by another fuzzy set. Zadeh [11] defined the concept of the inclusion measure of fuzzy set, and proposed a crisp relationship of the inclusion measure, which is either been included or been not included. Obviously this definition violates the characteristic of fuzzy set theory. After that many axiomatic definitions of inclusion measure of fuzzy sets were given in the literature. These

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 395–404.
springerlink.com

definitions revealed some essential properties of inclusion measure. Zadeh [11] also proposed the concept of fuzzy entropy which is used to describe the degree of fuzziness of fuzzy sets. Then some researchers investigated it with the different ways. For example, Burillo [2], Szmidt [7], Zhang [14], etc. gave different axiomatic definitions of entropy of intuitionistic fuzzy sets, respectively. However, axiomatic definitions of the entropy of intuitionistic fuzzy sets proposed by Szmidt [7] and of the interval-valued fuzzy sets proposed by Zeng [12] essentially are the same.

Zeng [12] extended the concepts of distance, similarity measure, inclusion measure and entropy to interval-valued fuzzy set theory, discussed the relationship between these numerical characteristics, and got some interesting conclusions. This article will extend these concepts to intuitionistic fuzzy set theory, research the relationship among them.

2 Preliminaries

In this section we briefly review basic knowledge of intuitionistic fuzzy sets and propose some new concepts of numerical characteristics of intuitionistic fuzzy sets.

Definition 1[1]. *Suppose X is the universe, then an intuitionistic fuzzy set A in X is given by: $A=\{<x,\mu_A(x),\nu_A(x)>| x\in X\}$, where the maps $\mu_A : X\to[0,1]$, $\nu_A : X\to[0,1]$ satisfy the condition: $0\le\mu_A(x)+\nu_A(x)\le 1$, $\forall x\in X$. $\mu_A(x)$ and $\nu_A(x)$ are called the membership and non-membership degrees of x to the intuitionistic fuzzy set A respectively.*

An intuitionistic fuzzy set A can be simply denoted: $A=<x,\mu_A(x),\nu_A(x)>$.Obviously, a classical fuzzy sets A is an intuitionistic fuzzy set: $A=<x,\mu_A(x),1-\mu_A(x)>$.

In this paper, we use S to denote the set of all intuitionistic fuzzy sets in X. Bustince and Burillo [4] showed that vague sets and intuitionistic fuzzy sets are the same extensions of fuzzy sets.

The inclusion relationship ($\subseteq$) and operations of complement (A^c), union ($A\cup B$), intersection ($A\cap B$) of intuitionistic fuzzy sets are defined as follows (see [12]):

Suppose A and B are intuitionistic fuzzy sets in X.

(1) We say that B includes A, denote $A\subseteq B$, if $\forall x\in X$, $\mu_A(x)\le\mu_B(x),\nu_A(x)\ge\nu_B(x)$.

(2) $A^c=<x,\nu_A(x),\mu_A(x)>$ is called the complement of A.

(3) $A\cup B=<x,\mu_A(x)\vee\mu_A(x),(1-\nu_A(x))\wedge(1-\nu_B(x))>$.

(4) $A\cap B=<x,\mu_A(x)\wedge\mu_A(x),(1-\nu_A(x))\vee(1-\nu_B(x))>$.

Next we introduce axiomatic definitions of distance, similarity measure, inclusion measure and entropy of intuitionistic fuzzy sets.

Definition 2[7]. *A real function E : S→[0,1] is called an entropy of S, if E satisfies the following conditions:*

(E1) $E(A)=0$ *iff* A *is a crisp sets;* (E2) $E(A)=1 \Leftrightarrow \mu_A(x)=\nu_A(x)$;

(E3) $E(A) \le E(B)$ *if* A *is less fuzzy than* B, *i.e.* $\mu_A(x) \le \mu_B(x)$ and $\nu_A(x) \ge \nu_B(x)$, *for* $\mu_B(x) \le \nu_B(x)$, $\mu_A(x) \ge \mu_B(x)$ and $\nu_A(x) \le \nu_B(x)$, for $\mu_B(x) \ge \nu_B(x)$; (E4) $E(A)=E(A^c)$.

According to the concept of similarity of interval-valued fuzzy sets from the literature [12], similarly, we can give the concepts of the similarity of intuitionistic fuzzy sets.

Definition 3. *A real function* S *: S×S→[0,1] is called a similarity measure of S, if* S *satisfies the following conditions:*

(S1) $S(A,A^c)=0$ *if* A *is a crisp set;* *(S2)* $S(A,B)=1 \Leftrightarrow A=B$;

(S3) $S(A,B)=S(B,A)$; *(S4) For all* $A,B,C \in S$, *if* $A \subseteq B \subseteq C$, *then*

$$S(A,C) \le S(A,B),\ S(A,C) \le S(B,C).$$

Obviously, the axiomatic definition of similarity measure of intuitionistic fuzzy sets is extended from fuzzy set theory. Particularly, if intuitionistic fuzzy sets A and B are fuzzy sets, then $S(A,B)$ is a similarity measure of fuzzy sets.

According to the concept of include measure of the interval-valued fuzzy sets from [3] and [12], we can propose the concepts of the include measure of intuitionistic fuzzy sets.

Definition 4. *A real function* I *: S×S→[0,1] is called an inclusion measure of intuitionistic fuzzy sets, if* I *satisfies the following conditions:*

(I1) $I(X,\varnothing)=0$; (I2) $I(A,B)=1 \Leftrightarrow A \subseteq B$;

(I3) For all $A,B,C \in S$, *if* $A \subseteq B \subseteq C$, *then* $I(C,A) \le I(B,A)$, $I(C,A) \le I(C,B)$.

According to Definition 3, we are able to define distances of intuitionistic fuzzy sets.

Definition 5. *A real function* d *: S×S→[0,1] is called a distance of intuitionistic fuzzy sets, if* d *satisfies the following conditions:*

(D1) $d(A, A^c) = 1$, *if* A *is a crisp set;* (D2) $d(A, B) = 0 \Leftrightarrow A = B$;

(D3) $d(A, B) = d(B, A)$;

(D4) *For all* $A, B, C \in S$, *if* $A \subseteq B \subseteq C$, *then*

$$d(A, C) \geq d(A, B),\ d(A, C) \geq d(A, B).$$

3 The Relationship among Four Numerical Characteristics

We construct a distance formula satisfies (D1-D4):

$$d_1(A, B) = \frac{1}{2n}\sum_{i=1}^{n}(\left|\mu_A(x_i) - \mu_B(x_i)\right| + \left|(1-\nu_A(x_i)) - (1-\nu_B(x_i))\right| \tag{1}$$

According to this distance, we can get the following inclusion measure:

$$I_1(A, B) = 1 - \frac{1}{2n}\sum_{i=1}^{n}(\left|\mu_A(x_i) - \mu_A(x_i) \wedge \mu_B(x_i)\right| + \left|(1-\nu_A(x_i)) - (1-\nu_A(x_i)) \wedge (1-\nu_B(x_i))\right| \tag{2}$$

Similarity, similarity measure can be got:

$$S_1(A, B) = 1 - \frac{1}{2n}\sum_{i=1}^{n}(\left|\mu_A(x_i) - \mu_B(x_i)\right| + \left|(1-\nu_A(x_i)) - (1-\nu_B(x_i))\right| = 1 - d_1(A, B)\,. \tag{3}$$

However, the entropy of S is difficult to be expressed directly by the distance of S, therefore we will use the following work to complete that the distance of S express the entropy of S.

According to the above formula, we have

Property 1

(1) $S_1(A, B) = S_1(A^c, B^c)$; (2) $S_1(A, B) = S_1(A \cap B, A \cup B)$;

(3) $S_1(A, A \cup B) = S_1(B, A \cap B) \Leftrightarrow S_1(A, A \cap B) = S_1(B, A \cup B)$.

Then, we discuss whether other similarity measures satisfy these three properties. Select several groups of similarity in [6] are given as following examples.

Example 1

$$S_2(A, B) = 1 - \frac{1}{\sqrt[p]{n}}\sqrt[p]{\sum_{i=1}^{n}|\, m_A(i) - m_B(i)\,|^p},\ m_C(i) = (\mu_C(x_i) + 1 - \nu_C(x_i))/2,\ A, B \in C.$$

Example 2

$$S_3(A,B)=1-\frac{1}{\sqrt[p]{n}}\sqrt[p]{\sum_{i=1}^{n}(\Phi_{tAB}(i)-\Phi_{fAB}(i))^p},\ \Phi_{tAB}(i)=|\mu_A(x_i)-\mu_B(x_i)|/2,$$

$$\Phi_{fAB}(i)=|(1-\nu_A(x_i))/2-(1-\nu_B(x_i))/2)|.$$

Example 3

$$S_4(A,B)=1-\frac{1}{\sqrt[p]{n}}\sqrt[p]{\sum_{i=1}^{n}(\Phi_{s1}(i)-\Phi_{s2}(i))^p},\Phi_{sn}(i)=|m_{An}(i)-m_{Bn}(i)|/2,1,2\in n,$$

$$m_{C1}(i)=(\mu_C(x_i)+m_C(i)),A,B\in C,\ m_{C2}(i)=(\mu_C(x_i)+1-\nu_C(x_i)),A,B\in C.$$

The distances correspond to above similarity measures are the following:

$$d_2=\frac{1}{\sqrt[p]{n}}\sqrt[p]{\sum_{i=1}^{n}|m_A(i)-m_B(i)|^p},d_3=\frac{1}{\sqrt[p]{n}}\sqrt[p]{\sum_{i=1}^{n}(\Phi_{tAB}(i)-\Phi_{fAB}(i))^p}$$

$$d_4=\frac{1}{\sqrt[p]{n}}\sqrt[p]{\sum_{i=1}^{n}(\Phi_{s1}(i)-\Phi_{s2}(i))^p},$$

It is not difficult to verify that S_3 and S_4 satisfy three conditions in Property 1. However, $S_2(A,B)\neq S(A\cap B,A\cup B)$, but $S_2(A,A\cup B)=S_2(B,A\cap B)$. This shows that not all the similarity measures satisfy these three properties, and Properties (2) and (3) can not be expressed by each other.

The above-mentioned similarity measures are of form $S(A,B)=1-d(A,B)$. Generally we can get following corollary.

Corollary 1. *The above-mentioned* $d_i(i=1,3,4)$ *satisfies the following properties:*

(1) $d_i(A,B)=d_i(A^c,B^c)$ (2) $d_i(A,B)=d_i(A\cap B,A\cup B)$

(3) $d_i(A,A\cup B)=d_i(B,A\cap B)\Leftrightarrow d_i(A,A\cap B)=d_i(B,A\cup B)$

$d_2(A,B)\neq d_2(A\cap B,A\cup B)$ *but* $d_2(A,A\cup B)=d_2(B,A\cap B)$.

Next, we discuss the relationships among the distance, the similarity measure, the inclusion measure, and the Entropy of intuitionistic fuzzy sets.

Theorem 1. *Suppose* d *and* S *are a distance and a similarity measure of intuitionistic fuzzy sets, respectively, then for* $A\in S$,

$$E'(A)=S(\mu_A(x),\ \nu_A(x)),\ E''(A)=1-d(\mu_A(x),\ \nu_A(x))$$

are entropies of intuitionistic fuzzy set A.

Proof. It is simple to prove E1, E2, E4. To save space, we just to prove E3.

(E3) For all $x_i \in X$, i =1, 2, . . . , n, if $\mu_B(x_i) \le \nu_B(x_i)$, i.e.,

$$\mu_A(x_i) \le \mu_B(x_i) \text{ and } \nu_A(x_i) \ge \nu_B(x_i).$$

Then $\mu_A(x_i) \le \mu_B(x_i) \le \nu_B(x_i) \le \nu_A(x_i)$. So by (S3) we have

$S(\mu_A(x_i), \nu_A(x_i)) \le S(\mu_A(x_i), \nu_B(x_i)) \le S(\mu_B(x_i), \nu_B(x_i))$ .i.e., $E'(A) \le E'(B)$.

Similarly, if $\mu_B(x_i) \ge \nu_B(x_i)$, i.e., $\mu_A(x_i) \ge \mu_B(x_i)$ and $\nu_A(x_i) \le \nu_B(x_i)$, then $\mu_A(x_i) \ge \mu_B(x_i) \ge \nu_B(x_i) \ge \nu_A(x_i)$. By (S3) we have,

$$S(\mu_A(x_i), \nu_A(x_i)) \le S(\mu_A(x_i), \nu_B(x_i)) \le S(\mu_B(x_i), \nu_B(x_i)).$$

Namely, we have $E'(A) \le E'(B)$.

If $S(\mu_A(x), \nu_A(x)) = 1 - d(\mu_A(x), \nu_A(x))$, $E'(A) = E''(A)$, $E''(A)$ is also an entropy of intuitionistic fuzzy set A.

Hence, we complete the proof of Theorem 1.

Corollary 2. *Suppose d and S are a distance and a similarity measure of intuitionistic fuzzy sets, respectively, for $A \in S$, then*

$$E'''(A) = S(A, A^c) \text{ and } E''''(A) = 1 - d(A, A^c)$$

are entropies of intuitionistic fuzzy set A.

The proof of this conclusion is similar to the proof of Theorem 1.

According to Definitions 4 and 5, we have the following conclusion.

Theorem 2. *Suppose that S is a similarity measure of intuitionistic fuzzy sets, then for $A, B \in S$, $I'(A, B) = S(A, A \cap B)$ and $I''(A, B) = S(B, A \cup B)$ are inclusion measures of intuitionistic fuzzy sets A and B.*

Proof. Firstly, we prove that $I'(A, B) = S(A, A \cap B)$ is an inclusion measure of intuitionistic fuzzy sets A and B.

It's simple to prove I1, I2, so we just to prove I3.

(I3) If $A \subseteq B \subseteq C$, then

$$I'(C, A) = S(C, C \cap A) = S(C, A),\ I'(B, A) = S(B, B \cap A) = S(B, A)$$

According to the definition of similarity measure of intuitionistic fuzzy sets, we have $I'(C,A) \le I'(B,A)$. Similarly, we have $I'(C,A) \le I'(C,B)$.

Similarly, we can prove that $I''(A,B) = S(B, A \cup B)$ is also an inclusion measure of intuitionistic fuzzy sets A and B.

Hence, we complete the proof of Theorem 2.

Theorem 3. *Suppose that I is an inclusion measure of intuitionistic fuzzy sets, for $A \in$ S, then $E(A) = I(A, A^c) \wedge I(A^c, A)$ is an entropy of intuitionistic fuzzy set A.*

Proof. It is simple to prove E1, E2, E4. To save space, we just to prove E3..

(E3) $\forall x_i \in X(\ i=1,2,\ldots,n\ )$, if $\mu_B(x_i) \le \nu_B(x_i)$, i.e., $\mu_A(x_i) \le \mu_B(x_i)$ and $\nu_A(x_i) \ge \nu_B(x_i)$. Then we have, $\mu_A(x_i) \le \mu_B(x_i) \le \nu_B(x_i) \le \nu_A(x_i)$,and

$$1\text{-}\nu_A(x_i) \le 1\text{-}\nu_B(x_i) \le 1\text{-}\mu_B(x_i) \le 1\text{-}\mu_A(x_i).$$

Hence we get $A \subseteq B \subseteq B^c \subseteq A^c$. Then by (I3) we have

$$I(A, A^c) \le I(B, A^c) \le I(B, B^c),\ I(A^c, A) \le I(A^c, B) \le I(B^c, B).$$

In other words, we have $E(A) = I(A, A^c) \wedge I(A^c, A) \le I(B, B^c) \wedge I(B^c, B) \le E(B)$.

Similarly, if $\mu_B(x_i) \ge \nu_B(x_i)$, we can obtain $E(A) \le E(B)$, too.

Hence, we complete the proof of Theorem 3.

In the highlight of Theorem 2, Corollary 3 and Theorem 3, we can get

$$E(A) = I(A, A^c) \wedge I(A^c, A) = S(A, A \cap A^c) \wedge S(A^c, A \cap A^c).$$

We know that the distance and the similarity measure of intuitionistic fuzzy sets are dual concepts, thus, we can use the distance to define the similarity measure of intuitionistic fuzzy sets. According to the relationship between the similarity measure and the distance of intuitionistic fuzzy sets which based on Hausdorff distance, we have

Theorem 4. *Given a real function f : [0,1]→[0,1]. If f is a strictly monotone decreasing function, and d is a distance of intuitionistic fuzzy sets, then for all A, $B \in$ S, $S(A,B) = \dfrac{f(d(A,B)) - f(1)}{f(0) - f(1)}$ is a similarity measure of intuitionistic fuzzy sets A and B.*

From D1-D4 we can easily prove that $S(A, B)$ satisfies S1-S4.

Now, our major problem is how to select a useful and reasonable f. In fact, the function $f(x) = 1 - x$ is such a simple and useful function.

It is well known that an exponential operation is highly useful in dealing with similarity relations, Shannon entropy in cluster analysis and other areas. We therefore choose the function $f(x) = e^{-x}$, then a similarity measure between A and B is defined as follows: $S_e(A, B) = \dfrac{e^{-d(A,B)} - e^{-1}}{1 - e^{-1}}$

On the other hand, we may choose $f(x) = 1/(1+x)$, then a similarity measure between A and B is defined as follows: $S_c(A, B) = \dfrac{1 - d(A, B)}{1 + d(A, B)}$.

If we choose $f(x) = 1 - d^p(A, B)$ $(1 \le p < \infty)$, then a similarity measure between

A and B is defined as: $S_p(A, B) = 1 - d^p(A, B)$.

Next, we discuss what properties $S_e(A, B)$, $S_c(A, B)$, $S_p(A, B)$ and their corresponding distance satisfy.

Theorem 5. *If* $S_e(A, B)$, $S_c(A, B)$ *and* $S_p(A, B)$ *satisfy Property 1(1-3), then the corresponding* $d(A, B)$ *satisfy Corollary 1(1-3).*

Proof. For $S_e(A, B) = \dfrac{e^{-d(A,B)} - e^{-1}}{1 - e^{-1}}$, $d(A, B) = -\ln((1 - e^{-1})S_e(A, B) + e^{-1})$ at axis $(d(A, B) \in [0,1],\ S_e(A, B) \in [0,1])$ with $S_e(A, B)$ is an isomorphism; For

$$S_c(A, B) = \frac{1 - d(A, B)}{1 + d(A, B)}, \quad d(A, B) = 2/(S_c(A, B) + 1) - 1$$

at axis $(d(A, B) \in [0,1],\ S_c(A, B) \in [0, 1])$ with $S_c(A, B)$ is an isomorphism; For $S_p(A, B) = 1 - d^p(A, B)$,

$$d(A,B)=\sqrt[p]{1-S_p(A,B)}$$

at axis ($d(A,B)\in[0,1], S_p(A,B)\in[0,1]$) with $S_p(A,B)$ is an isomorphism.

The four numerical characteristics of intuitionistic fuzzy sets include distance, similarity measure, inclusion measure and entropy, above theorems and corollaries show that E can be expressed by I or S, I can be expressed by S, and S and d can be expressed by each other as Fig.1.

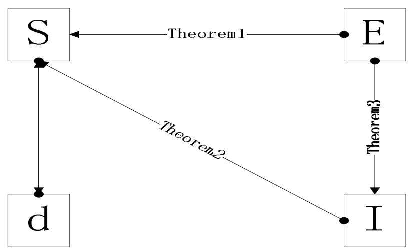


Fig. 1. Relation figure of four numerical characteristics

As Fig. 1, in order to achieve the four numerical characteristics can be expressed each other, we need a formula which can express S by E.

For the problem about S be expressed by E, by Theorem 2 in [13], we define A, $B\in$ S for all $x\in X$, define $T(A,B)$:

$$\mu_{T(A,B)}=(1+\min(|\mu_A(x)-\mu_B(x)|,|\nu_A(x)-\nu_B(x)|))/2,$$

$$\nu_{T(A,B)}=(1-\max(|\mu_A(x)-\mu_B(x)|,\ |\nu_A(x)-\nu_B(x)|))/2,$$

$T(A,B)\in$ S is obvious, then we have the following theorem.

Theorem 6. *Suppose* E *is an entropy of intuitionistic fuzzy sets, for* $A,B\in$ *S, then* $S(A,B)=E(T(A,B))$ *is a similarity measure of intuitionistic fuzzy sets* A *and* B.

This proof is similar with that of Theorem 2 in [13].

4 Conclusion

Considering the importance of numerical characteristics such as similarity measures, inclusion measures, entropies and distances of intuitionistic fuzzy sets, in this paper, we introduced an axiomatic definition of distances of intuitionistic fuzzy sets base on the axiomatic definition of the similarity measures. Our results show that the similarity measures, inclusion measures and entropies of intuitionistic fuzzy sets can be expressed by a distance of intuitionistic fuzzy sets. Furthermore, our results show that any one of the four numerical characteristics of intuitionistic fuzzy sets can be expressed by any other.

We believe that based on the obtained results in this paper, four numerical characteristics can be applied to more fields such as pattern recognition, image processing, approximate reasoning, fuzzy control, and so on.

References

1. Atanassov, K.: Intuitionistic fuzzy sets. Fuzzy Sets and Systems 20, 87–96 (1986)
2. Burillo, P., Bustince, H.: Entropy on intuitionistic fuzzy sets and on interval-valued fuzzy sets. Fuzzy Sets and Systems 78, 305–316 (1996)
3. Bustince, H.: Indicator of inclusion grade for interval-valued fuzzy sets. Application to approximate reasoning based on interval-valued fuzzy sets. International Journal of Approximate Reasoning 23, 137–209 (2000)
4. Bustince, H., Burillo, P.: Vague sets are intuitionistic fuzzy sets. Fuzzy Sets and Systems 79, 403–405 (1996)
5. Deschrijver, G., Kerre, E.: On the relationship between some extensions of fuzzy set theory. Fuzzy Sets and Systems 133, 227–235 (2003)
6. Hung, W.-L., Yang, M.-S.: Similarity measures of intuitionistic fuzzy sets based on Hausdorff distance. Pattern Recognition Letters 25, 1603–1611 (2004)
7. Szmidt, E., Kacprzyk, J.: Entropy for intuitionstic fuzzy sets. Fuzzy Sets and Systems 118, 467–477 (2001)
8. Wang, P.Z.: Fuzzy Sets and Its Applications. Shanghai Science and Technology Press (1983)
9. Wang, W., Xin, X.: Distance measure between intuitionistic fuzzy sets. Pattern Recognition Letters 26, 2063–2069 (2005)
10. Yu, T., Winslett, M., Seamons, K.E.: Interoperable strategies in automated trust negotiation. In: Proceedings of the 8th ACM Conference on Computer and Communications Security, pp. 146–155. ACM Press, New York (2001)
11. Zadeh, L.A.: Fuzzy sets. Information and control 8, 338–353 (1965)
12. Zeng, W., Guo, P.: Normalized distance, similarity measure, inclusion measure and entropy of interval-valued fuzzy sets and their relationship. Information Sciences 178, 1334–1342 (2008)
13. Zeng, W., Li, H.: Relationship between similarity measure and entropy of interval valued fuzzy sets. Fuzzy Sets and Systems 157, 1477–1484 (2006)
14. Zhang, H., Zhang, W., Mei, C.: Entropy of Interval-valued fuzzy sets based on distance and its relationship with similarity measure. Knowledge-Based Systems 22, 449–454 (2009)

The Study of Time Division Scheduling Mechanisms for Zigbee Cluster-Tree Wireless Sensor Networks

Jian-Ming Chen[1], Yan Xu[1], Rui Chang[1], and Guo-Xiu Peng[2]

[1] North China Institute of Water Conservancy and Hydroelectric Power, Zhengzhou, 450011, China
e-mail: {cjm,dlxuyan,changrui}@ncwu.edu.cn
[2] Beijing Shougang International Engineering Technology Co., Ltd, Beijing, 100043, China
e-mail: pengguoxiu@163.com

Abstract. To make a synchronized multi-hop cluster-tree network suitable for QoS support in WSNs, this paper unveils the applications of cluster-tree topology and proposes scheduling mechanisms approach based on time division. In this approach, time is divided into beacon frames and superframe duration that can be sent during the inactive period of its neighbor coordinators. Each coordinator transmits its beacon frame at the starting time Beacon_Tx_Offset,which is required to be different from that of its neighbor coordinators' and of its father coordinators'. This approach requires that when waken up, a coordinator should be active, so is its father node. The feasibility of this proposal has been exactly demonstrated through an experimental test platform based on an implementation of IEEE 802.15.4/Zigbee protocols built by ourselves.

Keywords: Zigbee, Cluster-tree, Superframe,Time Division, Synchronization.

1 Introduction

15.4b Workgroup [1] has been trying to improve IEEE 802.15.4 ,and to avoid beacon frame's collision , it has put forward some basic methods to be discussed, which may be adopted in the coming standard expansion. The first method is beacons' single-period method, in which we can set a time window at the beginning of the superframe that can be transmitted as a beacon frame. The second method is based on time division, in which the beacon frame of a special cluster is suggested to be transmitted when the other clusters are inactive. However, how to avoid beacon frames' collision is not presented.

2 Problem Discription

In cluster-tree networks of zigbee, beacon frame is used to synchronize every cluster, so if one node is within two ZCs communication range, collision of beacon

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 405–411.
springerlink.com

frames sent by the two ZCs may occur, which can result in failure of synchronization between this node and its father node.

Fig. 1 shows us a IEEE 802.15.4/Zigbee network containing N coordinators $\{ZR_i = (SD_i, BI_i)\}_{1\le i\le N}$, which can produce periodic beacon frames according to SO_i and BO_i. SD_i and BI_i stands for ZR_i's superframe duration and beacon interval respectively. The question is how to organize the beacon frames of the different coordinators to avoid their collision with other beacons or data frames by the method of time division. Obviously, we can transfer beacon frames by a continuous sequence, with which collisions between both direct and indirect neighbor coordinators can be avoided. Furthermore, beacon frames mustn't be transmitted in any superframe duration of other coordinators. Because each SD begins from beacon frame, beacon frame's scheduling returns to superframe scheduling. This problem is about non-priority scheduling of a group of periodic tasks, the time taken by which is individual to superframe continuing time, and its period is equal to BI. Accordingly, beacon is used to divide superframe duration when superframe schedule is running.

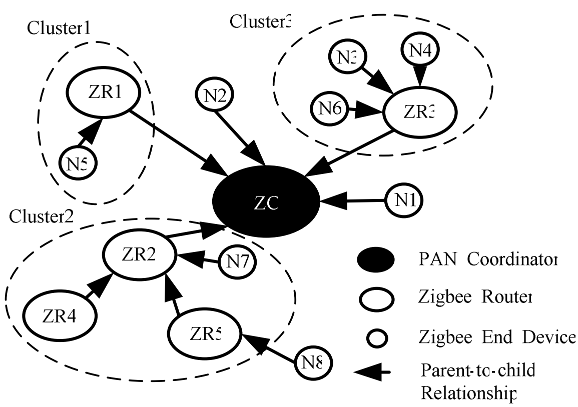


Fig. 1. Cluster-tree topological model

Two cases of superframe scheduling will be discussed later. Firstly, cases of constant SD (BOs may be different with each other); secondly, we extend conclusions of constant SD to cases of different SD.

We will discuss the two cases of superframe scheduling. Firstly, BOs are different; secondly, we can extend the conclusion to cases with different superframe duration.

3 Beacon Frame Scheduling Mechanisms for the Time Division Approach

3.1 Superframe Duration Scheduling algorithm for the Time Division Approach

In case of equal superframe durations, the superframe scheduling problem is somewhat similar to the pinwheel scheduling problem presented in [2-4]. The pinwheel problem consists in finding for a set of positive integer a cyclic schedule of indices $j \in (1,2,\cdots,n)$ such that there is at least one index j within any interval of a_j slots. By analogy to our problem, given a set of beacon intervals $A = (BI_1,\cdots,BI_N)$, the problem is to find a cyclic schedule of superframe durations such that there is at least one SD_i in each BI_i. In addition to the pinwheel problem, the distance between two consecutive instances of SD_i must be equal to BI_i. In this paper, we propose a general result for the scheduling problem for different and equal superframe durations.

The definition of a set C_M is:

$$C_M = \begin{cases} A \middle| A = \{a_1,\cdots,a_n\}, \\ when i < j\ ,\ a_j \text{ is exactly divisible by } a_i \ \sum 1/a_i \leq 1. \end{cases} \tag{1}$$

If a circular schedule exists, the least period will be the least common multiple of all the set's members, which can be mathematically described as: $LCM(a_1,a_2,\cdots,a_n) = \max\limits_{1\leq i\leq N}(a_i)$.

Proof. The proof is made by contradiction. Assume that a cyclic schedule exists for an instance $A \in \square_M$ of the pinwheel problem. Since $\forall i < j \Rightarrow a_i$ divide a_j, the $\forall i < j$ it exists an integer k_{ij} such that $a_j = k_{ij} \cdot a_i$ (harmonic integers). Then, we have $LCM(a_1,a_2,\cdots,a_n) = \max\limits_{1\leq i\leq N}(a_i)$.

Assume that the minimum cycle length is different from $LCM(a_1,a_2,\cdots,a_n)$. Then, since $LCM(a_1,a_2,\cdots,a_n)$ is not a cycle length, it exists a time slot n that contains a_i such that the $(n+LCM(a_1,a_2,\cdots,a_n))^{th}$ time slot does not contain a_i. Since $LCM(a_1,a_2,\cdots,a_n)$ is a multiple of a_i, it directly implies that the set is not schedulable, which is absurd.

Therefore, to perform the scheduling operation of inconstant SD and BI superframe sequence, we put forward Superframe Duration Scheduling(SDS) algorithm, according to which we can produce a schedule list of schedulable superframe sequence. The algorithm is also applicable to constant period superframe sequence.

The necessary condition of a network with N coordinators $\{ZR_i = (SD_i, BI_i)\}_{1\le i\le N}$ whose SD_i are different is

$$\sum_{i=1}^{N} DC_i = \sum_{i=1}^{N} \frac{SD_i}{BI_i} \le 1 \tag{2}$$

Define superframe duration $\overline{BI}_{maj} = LCM\left(2^{bo_1}, 2^{bo_2}, \cdots, 2^{bo_n}\right) = \max\limits_{1\le i\le N}\left(2^{BO_i}\right)$, known as major cycle, while the least BI is known as minor cycle.

The idea of the SDS algorithm is the following:

(1) Describe the set of beacon frame intervals as $A = \{2^{BO_i}\}_{1\le i\le N}$.

(2) The minimum beacon interval $\overline{BI}_{\min} = 2^{BO_{\min}}$ *is calculated.*

(3) Sequence the members of A in the increasing order of BO_i.

(4) If $(\overline{BI}_i = \overline{BI}_j)$, *then*

(5) If $(\overline{SD}_i \ge \overline{SD}_j)$, put $\overline{BI}_i$ precede to $\overline{BI}_j$;

(6) Or else put $\overline{BI}_j$ precede to $\overline{BI}_i$;

(7) Divide $\overline{BI}_{maj}$ by $\min(SD_i)_{1<i<N}$ into n slot time;

(8) For (each element i in A), do {

(9) Search the first retrievable slot time whose length is SD_i;

(10) Write(i) write down the retrievable slot time;

(11) Repeat.

(12) If (write (i)=*false*).

(13) Then return (“*not schedulable*”).

(14) Until(reaching the end of the major cycle)}.

(15) Return(“schedule terminates”).

3.2 Superframe Duration Scheduling with Coordinator Grouping

Here, we extend time division method to optimum superframe scheduling algorithm of large-scale network. Results show that some coordinators are so far away from each other that their transmitting ranges don’t overlap, they can transmit beacons simultaneously without direct or indirect beacon frame collision.

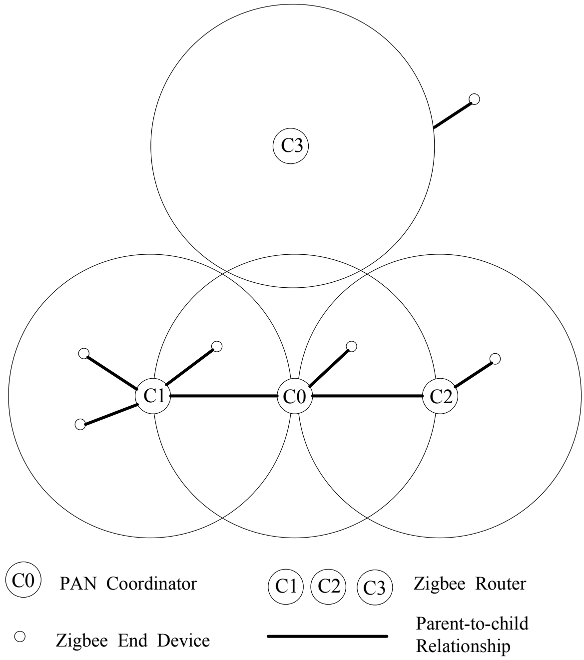


Fig. 2. The geographic distribution of the nodes in the network

To illustrate this method intuitively, we take an example as Fig. 2, which has shown us a distributive network, whose parameters are listed in Table 1 (SD is a unit of time).

Table 1. Example of PAN configuration

Coordinator	SO	BO	SD	BI
C0	0	1	1	2
C1	0	1	1	2
C2	0	1	1	2
C3	0		1	2

Fig. 2 demonstrates that collisions among beacons from C0, C1 and C2 occur because of the transmitting range of C0 and that of C1 and C2 overlap. According to the above formula, the four coordinators cannot perform superframe duration schedule because the total duty cycle is bigger than 1 $(0.5+0.5+0.5=1.5>1)$.

However, because C1, C2 and C3 is neither direct nor indirect neighbour (their transmitting range doesn't overlap), they can send beacon frame simultaneously.

Thus, we allow C1, C2 and C3 send beacon frame simultaneously, C0 follows behind. In this case, beacon frame collision will not occur and this group of coordinators becomes schedulable, as shown in Fig. 3.

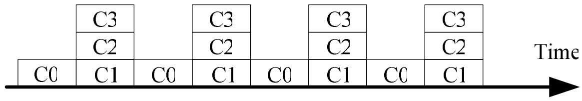


Fig. 3. Superframe duration scheduling with coordinator grouping

The common method of node group's sending beacon frame simultaneously is described as follows: supposed each coordinator's transmission range is a circle whose radius is r. The nonoverlapping of two coordinators means the distance between them is at least 2·r, and the simultaneously sending beacon frames is allowed by the two coordinators. When the vertex indicates coordinator, arris indicates the line of the length more than 2·r used to connect two coordinators, we need to consider vertex coloration in graphics [5-6], whose algorithm can be realized in PAN coordinator. Proposed the position of all the coordinators in the network is known, and they are split up into groups, the information of which is returned to nodes. After vertex coloration is carried out, the coordinators with the same color and in the same group can send beacon frame at the same time as all the coordinators in the same group.

The advantage of the grouping policy is: a schedule for the group of coordinators whose duty cycle is greater than 1(as in the above example).

4 Implementation Introduction

Factually, superframe scheduling algorithm (without coordinators' grouping) can be easily realized by modifying IEEE 802.15.4 somewhat. When a new coordinator has joined in the network, it sends superframe structure regulations(BO and SO) in single-hop form to PAN coordinator, which call the above algorithm to analyze the whole coordinator group's schedule ability(containing the new one's). If the algorithm can produce a valid schedule list, the new coordinator will be allowed to send beacon frame. Simultaneously, the new schedule list is returned to all the nodes by beacon frame. Then, all the coordinators update their offset, referring to which beacon frame is sent. What we should pay attention to is a coordinator's offset is decided in reference to its father coordinator's beacon transmission time. And if a valid is not produced by the algorithm, the corresponding coordinator will not be allowed to send beacon frame.

5 Conclusion

In this paper, according to improve IEEE 802.15.4/Zigbee cluster-tree standard, we propose superframe scheduling algorithm, which arrange superframe time sequence of different coordinators judging by SO and BO in non-overlap mode. At the same time, we point out the method could be improved by dividing coordinators into groups, which would increase complexity when carried out.

An important step in this method is understanding the extension structure's complexity of cluster tree in IEEE 802.15.4/Zigbee WPANs and finding a method of expansion. Presently, these methods are run and tested on the experiment platform, and the protocol stack's basic functions of IEEE 802.15.4/Zigbee have been realized [7].

Acknowledgements. This project is supported by the Scientific and Technological Research of Henan Province（082102210112）.

References

1. The Institute of Electrical and Electronics Engineers. IEEE 802.15 WPAN™ Task Group 4b (TG4b) [DB/OL] (February 2008), http://grouper.ieee.org/groups/802/15/pub/TG4b.html
2. Holte, R., Mok, A., Rosier, L., Tulchinsky, I., Varvel, D.: The pinwheel: A Real-Time Scheduling Problem. In: Proceedings of the 22nd Hawaii International Conference on System Science, Hawaii, pp. 693–702 (1989)
3. Chen, H., Xiong, G.: An On-Line Scheduler over Hard Real-Time Communication System. Journal of Electronic Science and Technology of China 1, 53–59 (2003)
4. Lin, H.-h., Hsueh, C.-W.: Power-Aware Real-Time Scheduling using Pinwheel Model and Profiling Technique. In: Proc. of the 11th International Conference on Real-Time and Embedded Computing Systems and Applications (RTCSA 2005, Best Paper Award), Hong Kong (August 2005)
5. Diestel, R.: Graph Theory, 3rd edn. Springer, New York (2005)
6. Xu, J., Qiang, X., Fang, G.: A DNA computer model for solving vertex coloring problem. Chinese Science Bulletin 20, 2541–2549 (2006)
7. Crossbow Technology, MICAz datasheet [DB/OL] (2008), http://www.xbow.com

Discovery of S-Rough Decision Law and Its Relation Metric

Jin-Ming Qiu and Chang-Jing Lu

Department of Mathematics and Computer Science,
Sanming University, Sanming Fujian 365004, China
e-mail: fjsmqjm@163.com, smlcj123@163.com

Abstract. For dynamic decision-making system law mining, by using two direction S-rough decision model, this paper presents the concepts of two direction S-rough decision law, and F -law decision relation metric; by using these concepts, this paper proposes characteristic theorem of F -rough decision law relation metric, dependence theorem of F -rough decision law relation metric, the principle of F -rough decision law discovery and gives the application of F -rough decision law in profit risk decision analysis.

Keywords: Function two direction S-rough sets, F -decision rough law, rough decision law relation metric, the discovery principle of F -rough decision law.

1 Introduction

Many decision-making systems, such as economic decision-making system, management and decision-making system are always interfered by the interior factors, or the exterior factors, which makes the decision-making factors in the decision-making sets change, getting more or less, and the dynamic change of the sets must cause the dynamic change of the decision-making. As a result, the sequence decision-making $u_0 = (u_0(1), u_0(2), \cdots, u_0(n))$, which is gotten by the decision-making system dealing with some sequence event $A_0=(a_0(1), a_0(2), \ldots, a_0(n))$ is changed to $u' = (u'(1), u'(2), \cdots, u'(n))$, moreover $u' \neq u_0 (\exists k, u'(k) \neq u_0(k), k \in \{1, 2, \cdots, n\}), u'$ deviates from the given u_0. This fact indicate that there is a relation between u' and u_0. Assume that the sequence decision-making $u_i = (u_i(1), u_i(2), \cdots, u_i(n))$, which is gotten by the decision-making system dealing with some sequence event $A_i= (a_i(1), a_i(2), \ldots, a_i(n))$ is a discrete law, then the sequence decision-making family gotten by the decision-making system dealing with sequence event set $A=\{A_1, A_2, \ldots, A_m\}$ make up of an uncertain decision-making discrete law set or discrete function set $Q^* = \{u_i \mid u_i = (u_i(1), u_i(2), \cdots, u_i(n)), i = 1, 2, \cdots, m\}$. The attributes are supplemented or deleted from the decision attribute set α is equal to that the decision factors in the decision factor set getting more or less. Obviously, if we treat the disturbance to the interior or the exterior factors of the

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 413–424.
springerlink.com

decision-making system as the decision factors deleting or supplementing to the decision factor set α^* of decision law set Q^*, and the decision law can compose to different decision law equivalence class according to the different characteristics of the decision factors or decision attributes, then this process can be abstracted to a phenomenon in function two direction S-rough sets [1-10]: the attribute set α of function equivalence class $[u]$ is under F - disturbance, then the elements in function equivalence class $[u]$ are supplemented or deleted. This fact is consistent to the characteristic of function two direction S-rough sets. Based on this fact, people wonder that if we can study the decision law of the dynamic decision-making system by using the law characteristic and the dynamic characteristic of function two direction S-rough sets [1-10], and to know the relationship among the decision laws, as well as using the relations to adjust and control the decision-making of the system.

The main results of this paper is that giving the decision law and its existence, $(f,\overline{f})$ -decision law generation, F -decision rough law generation and its charac-ter-ristics, $(f,\overline{f})$ -decision law relation metric theorem, and F -rough decision law relation metric theorem , discussing the discovery principle of F -rough decision law and giving the applications.

2 Two Direction S-Rough Decision

Let two direction S-sets X^* be a decision factors (target) set of a decision event a, $D=\{d_1,d_2,\cdots,d_n\}$ be decision-making countermeasure set. $u_j^{(x)*}$ is decision on countermeasure d_j of $(X)^* = (R,F)_\circ(X^*)$, or called the lower decision of countermeasure d_j; $u_j^{(x)*}\in[0,1]$, moreover

$$u_j^{(x)*} = 1/(1+\left\{\sum_{i=1}^{\sigma}(w_i^{(x)*}(g_i^{(x)*}-r_{ij}^{(x)*}))^p \Big/ \sum_{i=1}^{\sigma}(w_i^{(x)*}(r_{ij}^{(x)*}-b_i^{(x)*}))^p\right\}^{2/p}) \tag{1}$$

Where $j=1,2,\cdots,n; r_{ij}^{(x)*}, g_i^{(x)*}, b_i^{(x)*}, w_i^{(x)*}$ is the target superior degree, maximum superior degree, minimum superior degree, target weigh and of the ith decision factor of d_j with respect to $(X)^*$, and $\sum_{i=1}^{\sigma} w_i^{(x)*}=1$.

By using (1), the lower decision set of event a is obtained, moreover

$$\{u_1^{(x)*},u_2^{(x)*},\cdots,u_n^{(x)*}\} \tag{2}$$

$u_j^{(y)*}$ is rough decision on countermeasure d_j of $(Y)^* = (R,F)^\circ(X^*)$, or called the upper decision of counter- measure d_j with respect to $(Y)^*$; $u_j^{(y)*}\in[0,1]$, moreover

$$u_j^{(y)*} = 1/(1+\left\{\sum_{i=1}^{\tau}(w_i^{(y)*}(g_i^{(y)*} - r_{ij}^{(y)*}))^p / \sum_{i=1}^{\tau}(w_i^{(y)*}(r_{ij}^{(y)*} - b_i^{(y)*}))^p\right\}^{2/p}) \tag{3}$$

By using (3), the upper decision set of event *a* is obtained, moreover

$$\{u_1^{(y)*}, u_2^{(y)*}, \cdots, u_n^{(y)*}\} \tag{4}$$

Assume that sorting sequence on (2), (4):

$$u_1^{(x)*} \le u_2^{(x)*} \le \cdots \le u_n^{(x)*} \tag{5}$$

$$u_1^{(y)*} \le u_2^{(y)*} \le \cdots \le u_n^{(y)*} \tag{6}$$

Two direction S-rough decision set of event *a* is obtained from (5),(6), moreover

$$\{(u_1^{(x)*}, u_1^{(y)*}), (u_2^{(x)*}, u_2^{(y)*}), \cdots (u_n^{(x)*}, u_n^{(y)*})\} \tag{7}$$

Where $(u_j^{(x)*}, u_j^{(y)*})$ is a rough decision of X^*. sequence (7) can be used in choosing decision-making.

The following result can be obtained easily.

Theorem 1. *If decision factors (target) sets $X_p^* \ne X_q^*$ of event a, then Rough decision of X_p^*, X_q^* fulfill*

$$(u_p^{(x)*}, u_p^{(y)*}) \ne (u_q^{(x)*}, u_q^{(y)*}) \tag{8}$$

Theorem 2. *If rough decision $(u_j^{(x)*}, u_j^{(y)*})$ of X^* fulfill*

$$u_j^{(x)*} = u_j^{(y)*}, \tag{9}$$

then

$$(R, \mathrm{F})_{\circ}(X^*) = (R, \mathrm{F})^{\circ}(X^*) \tag{10}$$

More concepts of function S-rough sets and S-rough decision can be found in [1-4].

Base on two direction S-rough decision, the discussion about F -rough decision law relation is given in section 3.

3 Decision Law Generation and *F* -Decision Law Relation Metric

According to section 2, two direction S-rough decision set of sequence event A_i= ($a_i(1)$, $a_i(2)$, …, $a_i(n)$) can be obtained, moreover

$$\{(u_{i1}^{(x)*}, u_{i1}^{(y)*}), (u_{i2}^{(x)*}, u_{i2}^{(y)*}), \cdots, (u_{in}^{(x)*}, u_{in}^{(y)*})\} \tag{11}$$

Where $(u_{ij}^{(x)*}, u_{ij}^{(y)*})$ is two direction S-rough decision of event a_{ij}, j=1, 2, **…**, n.

Definition 1. $(\underline{u}_i, \overline{u}_i)$ *is called two direction S-rough decision law that is generated by the sequence event A_i= (a_i(1), a_i(2), …, a_i(n)), for short two direction S-rough decision law of A_i, moreover*

$$(\underline{u}_i, \overline{u}_i) = ((u_{i1}^{(x)*} u_{i1}^{(y)*}), (u_{i2}^{(x)*} u_{i2}^{(y)*}), \cdots, (u_{in}^{(x)*}, u_{in}^{(y)*})) \tag{12}$$

Where $\underline{u}_i, \overline{u}_i$ are the lower decision law and the upper decision law of $(\underline{u}_i, \overline{u}_i)$, respectively; or called the lower decision law and the upper decision law of A_i, respectively. $(u_{ij}^{(x)*}, u_{ij}^{(y)*})$ is two direction S-rough decision of event a_{ij}, j=1, 2, **…**, n.

According to Definition 1, it follows: event set A**=**{A_1**,** A_2**, …,** A_m} generates the two direction S-rough decision set $Q^* = \{(\underline{u}_i, \overline{u}_i) \mid i = 1, 2, \cdots, m\}$, the two direction S-lower decision law set $\underline{Q}^* = \{\underline{u}_i \mid i = 1, 2, \cdots, m\}$, and the two direction S-upper decision law set $\overline{Q}^* = \{\overline{u}_i \mid j = 1, 2, \cdots, m\}$. Obviously, With a common attribute set of the whole decision on law constitutes decision law equivalence class $[u] = \{u_1, u_2, \cdots, u_m\}$; And the attributes are supplemented or deleted from the decision attribute set α is equal to that the decision factors in the decision factor set X^* getting more or less. When the attributes are supplemented or deleted from the decision attribute set α, the decision law equivalence class $[u]'$ with changed attribute set α can be obtained.

Definition 2. *Decision law* $[u]^{(f,\overline{f})} = \{u_1, u_2, \cdots, u_\lambda\}$ *is called the* $(f, \overline{f})$ *-decision law of decision law* $[u] = \{u_1, u_2, \cdots, u_m\}$; *if the attribute set* $\alpha^{(f,\overline{f})}$ *of* $[u]^{(f,\overline{f})}$ *and the attribute set* α *of* $[u]$ *satisfy*

$$\alpha^{(f,\overline{f})} = \alpha \cup \{\alpha_t \mid f(\beta_t) = \alpha_t \in \alpha, \beta_t \overline{\in} \alpha\} - \{\alpha_s \mid \alpha_s \in \alpha, \overline{f}(\alpha_s) = \beta_s \overline{\in} \alpha\} \tag{13}$$

$u^{(f,\overline{f})} = (x^{(f,\overline{f})}(1), x^{(f,\overline{f})}(2), \cdots, x^{(f,\overline{f})}(n))$ *and* $u = (x(1), x(2), \cdots, x(n))$ *are called the composition of* $[u]^{(f,\overline{f})}$ *and* $[u]$ *respectively, if*

$$\forall x^{(f,\overline{f})}(k) \in u^{(f,\overline{f})},\ x^{(f,\overline{f})}(k) = \sum_{i=1}^{\lambda} x_i(k) \tag{14}$$

$$\forall x(k) \in u, x(k) = \sum_{j=1}^{m} x_j(k),\ k = 1, 2, \cdots, n. \tag{15}$$

Where

$\forall u_i \in [u]^{(f,\overline{f})}, u_i = (x_i(1), x_i(2), \cdots, x_i(n)), i = 1, 2, \cdots, \lambda; \forall u_j \in [u], u_j = (x_j(1), x_j(2), \cdots, x_j(n)), j = 1, 2, \cdots, m, x_j(t) \in \mathbb{R}, t = 1, 2, \cdots, n; \beta_t \overline{\in} \alpha, f(\beta_t) = \alpha_t \in \alpha, \alpha_s \in \alpha, \overline{f}(\alpha_s) = \beta_s \overline{\in} \alpha, \beta_s, \beta_t \in V, \alpha = \{\alpha_1, \alpha_2, \cdots, \alpha_r\}$, V is the attribute universe.

Definition 3. *The decision law pair which is composed of* $[u]_-$ *and* $[u]^-$

$$([u]_-,[u]^-) \tag{16}$$

is called rough decision law generated from function two direction S-rough sets, rough decision law $([u]_-,[u]^-)$ *for short.*

Definition 4. *The* F *- decision law pair which is composed of* $[u]_F$ *and* $[u]^F$

$$([u]_F,[u]^F) \tag{17}$$

is called F *-rough decision law generated from* $([u]_-,[u]^-)$, F *-rough decision law* $([u]_F,[u]^F)$ *for short.*

Where $[u]_F$ is the F - decision law of $[u]_-$, and $[u]_F = \cup_i [u]_i^{(f,\bar{f})} = (R,F)_\circ(Q^*)_F$, $[u]_- = \cup_i [u]_i = (R,F)_\circ(Q^*)$; $[u]^F$ is the F - decision law of $[u]^-$, and $[u]^F = \cup_j [u]_j^{(f,\bar{f})} = (R,F)^\circ(Q^*)^F$, $[u]^- = \cup_j [u]_j = (R,F)^\circ(Q^*)$ [1-6].

Definition 5. *Let* $[u]_i^{(f,\bar{f})}$ *be the* $(f,\overline{f})$ *-decision law of* $[u]_0$, $\mu(u_i^{(f,\bar{f})}(k),u_0(k))$ *is called the decision law relation coefficient of* $[u]_i^{(f,\bar{f})}$ *with respect to* $[u]_0$ *at point* k, *if*

$$\mu(u_i^{(f,\bar{f})}(k),u_0(k)) = \rho \max_{i,k} \Delta_{0,i}(k) / \Delta_{0,i}(k) + \rho \max_{i,k} \Delta_{0,i}(k) \tag{18}$$

where $k = 1,2,\cdots,n-1$.

$$\Delta_{0,i}(k) = S_{0,i}(k) + \overline{S}_{0,i}(k) \ , \ S_{0,i}(k) = | P_{0,i}(k+1) - P_{0,i}(k) | \tag{19}$$

$$\overline{S}_{0,i}(k) = \begin{cases} (1/2)\cdot(| P_{0,i}(k+1) + P_{0,i}(k) |), P_{0,i}(k+1)P_{0,i}(k) \geq 0 \\ | P_{0,i}(k+1) |^2 + | P_{0,i}(k) |^2 / 2(| P_{0,i}(k+1) - P_{0,i}(k) |), P_{0,i}(k+1)P_{0,i}(k) < 0 \end{cases} \tag{20}$$

$$P_{0,i}(k) = x_0(k) - x_i(k) \tag{21}$$

$\rho \in (0,1]$ is the detached coefficient of $[u]_i^{(f,\bar{f})}$ about $[u]_0$, $u_i^{(f,\bar{f})} = (x_i^{(f,\bar{f})}(1), x_i^{(f,\bar{f})}(2),\cdots, x_i^{(f,\bar{f})}(n))$, $u_0 = (x_0(1), x_0(2), \cdots x_0(n))$, $x_i^{(f,\bar{f})}(k), x_0(k) \in \mathbb{R}$.

Definition 6. $\gamma(u_i^{(f,\bar{f})},u_0)$ *is called the decision law relation metric of decision law* $[u]_i^{(f,\bar{f})}$ *with respect to decision law* $[u]_0$, *if*

$$\gamma(u_i^{(f,\bar{f})},u_0) = (1/n)\cdot\sum_{k=1}^{n-1}\mu(u_i^{(f,\bar{f})}(k),u_0(k)) \tag{22}$$

Definition 7. *Let* $([u]_F, [u]^F)$ *be the* F*-decision law of* $([u]^-, [u]_-), (\mu(u_F(k), u_-(k)), \mu(u^F(k), u^-(k)))$ *and* $(\gamma(u_F, u_-), \gamma(u^F, u^-))$ *are called the decision law relation coefficient and the decision law relation metric of* $([u]_F, [u]^F)$ *about* $([u]_-, [u]^-)$ *at point* k*, respectively.*

Where (u_F, u^F) is the composition of $([u]_F, [u]^F)$; the structure of $\mu(u_F(k), u_-(k))$, $\mu(u^F(k), u^-(k))$, $\gamma(u_F, u_-)$ and $\gamma(u^F, u^-)$ are the same as the forms of (18) and (22), respectively. By definitions 1-7, the following fact can be obtained.

Theorem 3. *(Characteristic theorem of* $(f, \bar{f})$*-decision law metric). If* $\gamma(u_i^{(f,\bar{f})}, u_0)$ *is the* $(f, \bar{f})$*-decision law relation metric of* $[u]_i^{(f,\bar{f})}$ *with respect to* $[u]_0$*, then*

1° *Regularity.* $0 < \gamma(u_i^{(f,\bar{f})}, u_0) \le 1$;

2° *Symmetry.* $\gamma(u_1^{(f,\bar{f})}, u_0) = \gamma(u_0, u_1^{(f,\bar{f})})$;

3° *Non-uniformity.* $\gamma(u_k^{(f,\bar{f})}, u_0) \ne \gamma(u_0, u_k^{(f,\bar{f})})$, $k > 2$.

Proof. 1° By (19), it follows $\Delta_{0,i}(k) \ge 0; \forall i, k$, by (22), there is $0 < \gamma(u_i^{(f,\bar{f})}, u_0) \le 1$. $\gamma(u_i^{(f,\bar{f})}, u_0) = 1 \Leftrightarrow \sum_{k=1}^{n-1} \mu(u_i^{(f,\bar{f})}(k), u_0(k)) = n-1$, $0 < \mu(u_i^{(f,\bar{f})}(k), u_0(k)) \le 1 \Leftrightarrow \mu(u_i^{(f,\bar{f})}(k), u_0(k)) = 1 \Leftrightarrow \Delta_{0,i}(k) = 0 \Leftrightarrow S_{0,i}(k) = 0$ and $\bar{S}_{0,i}(k) = 0 \Leftrightarrow u_i^{(f,\bar{f})}(k) = u_0(k) \Leftrightarrow u_i^{(f,\bar{f})} = u_0$.

2° If $u = \{u_1^{(f,\bar{f})}, u_0\}$, then there must be $\max_{i,k} \Delta_{0,i}(k) = \max_k \Delta_{0,1}(k) = \max_k \Delta_{1,0}(k) = \max_{i,k} \Delta_{i,0}(k)$ in (4), so $\gamma(u_1^{(f,\bar{f})}, u_0) = \gamma(u_0, u_1^{(f,\bar{f})})$. 3° If $u^{(f,\bar{f})} = \{u_j^{(f,\bar{f})} \mid j = 1, 2, \cdots, m; \quad m > 2\}$, since $\max_{j,k} \Delta_{0,j}(k) \ne \max_{j,k} \Delta_{j,0}(k)$, then there must be 3°.

Theorem 4. *(Characteristic theorem of* F*-decision law relation metric). If* $(\gamma(u_i^F, u^-), \gamma(u_{F,i}, u_-))$ *is the* F*-decision law relation metric of* $([u]_i^F, [u]_{F,i})$ *with respect to* $([u]^-, [u]_-)$*, then*

1° *Regularity.* $0 < \gamma(u_i^F, u^-) \le 1$, $0 < \gamma(u_{F,i}, u_-) \le 1$;

2° *Symmetry.* $\gamma(u_1^F, u^-) = \gamma(u^-, u_1^F)$, $\gamma(u_{F,1}, u_-) = \gamma(u_-, u_{F,1})$

3° *Non-uniformity.* $\gamma(u_k^F, u^-) \ne \gamma(u^-, u_k^F)$, $\gamma(u_{F,k}, u_-) \ne \gamma(u_-, u_{F,k})$, $k > 2$.

Theorem 4 can be proved in a similar way as shown theorem 3 and it is omitted.

4 Relation Metric and the Discovery of F -Rough Decision Laws

Definition 8. *The band composed of rough decision law* $([u]_{-},[u]^{-})$ *is called rough decision law band, moreover*

$$\mathrm{BAND}\{D_{-},D^{-}\} \tag{23}$$

D_{-},D^{-} *are called the lower boundary and upper boundary of* $\mathrm{BAND}\{D_{-},D^{-}\}$ *respectively .*

The band composed of F -rough decision laws $([u]_{F},[u]^{F})$ is called F -rough decision laws band, moreover

$$\mathrm{BAND}\{D_{F},D^{F}\} \tag{24}$$

D_{F},D^{F} are called the lower boundary and upper boundary of $\mathrm{BAND}\{D_{F},D^{F}\}$ respectively. Where $D_{-}=u_{-},D^{-}=u^{-}$; $D_{F}=u_{F},D^{F}=u^{F}$.

Definition 9. *The* F *-decision law relation coefficient pair composed of* $\mu(u_{F}(k),u_{-}(k))$ *and* $\mu(u^{F}(k),u^{-}(k))$ *,*

$$(\mu(u_{F}(k),u_{-}(k)),\mu(u^{F}(k),u^{-}(k))) \tag{25}$$

is called the F *-rough decision law relation coefficients of* $([u]_{F},[u]^{F})$ *with respect to* $([u]_{-},[u]^{-})$ *at point* k *.*

Definition 10. *The* F *- rough decision law relation metric pair composed of* $\gamma(u_{F},u_{-})$ *and* $\gamma(u^{F},u^{-})$ *,*

$$(\gamma(u_{F},u_{-}),\ \gamma(u^{F},u^{-})) \tag{26}$$

is called the F *-rough decision law relation metric of* $([u]_{F},[u]^{F})$ *with respect to* $([u]_{-},[u]^{-})$.

By definitions 6-10, the following fact can be obtained.

Theorem 7. *(Dependence theorem of* F *-rough decision law relation metric). Let* $([u]_{F,i},[u]_{i}^{F})$ $([u]_{F,j},[u]_{j}^{F})$ *be* F *-rough decision law respectively;* $(\alpha_{-},\alpha^{-}),(\alpha_{F,i},\alpha_{i}^{F})$ *and* $(\alpha_{F,j},\alpha_{j}^{F})$ *be the attribute sets of* $([u]_{F,i},[u]_{i}^{F})$ *,* $([u]_{-},[u]^{-})$, $([u]_{F,i},[u]_{i}^{F})$ *and* $([u]_{F,j},[u]_{j}^{F})$ *respectively,*

if
$$(\alpha_{-},\alpha^{-})\Rightarrow(\alpha_{F,i},\alpha_{i}^{F})\Rightarrow(\alpha_{F,j},\alpha_{j}^{F}) \tag{27}$$

then
$$\gamma(u_{F,j},u_{-})\le\gamma(u_{F,i},u_{-}),\ \gamma(u_{j}^{F},u^{-})\le\gamma(u_{i}^{F},u^{-}) \tag{28}$$

Where ($[u]_-,[u]^-$) is rough decision law; u_-,u^- are the composition of $[u]_-$ and $[u]^-$ respectively; and $\alpha_{F,i} \Rightarrow \alpha_{F,j}, \alpha_i^F \Rightarrow \alpha_j^F$, are written as $(\alpha_{F,i}, \alpha_i^F) \Rightarrow (\alpha_{F,j}, \alpha_j^F)$.

Proof. Since $\alpha_- \Rightarrow \alpha_{F,i} \Rightarrow \alpha_{F,j}$, namely $\alpha_- \subseteq \alpha_{F,i} \subseteq \alpha_{F,j}$, $[u]_{F,j} \subseteq [u]_{F,i} \subseteq [u]_-$ Assume that $u_-, u_{F,i}, u_{F,j}$ are the composition of $[u]_-$, $[u]_{F,i}$ and $[u]_{F,j}$ respectively, obviously, there is $u_{F,j} \le u_{F,i} \le u_-, \forall\lambda \in \{1,2,\cdots,n\}, x_{F,j}(\lambda) \le x_{F,i}(\lambda) \le x_-(\lambda)$; $x_-(\lambda) \in u_-, x_{F,i}(\lambda) \in u_{F,j}$. namely By (18)-(22), it follows $\gamma(u_{F,j},u_-) \le \gamma(u_{F,i},u_-)$, namely where u_- is the composition of $[u]_-$, and $u_- = (x_-(1), x_-(2), \cdots, x_-(n)), x_-(\lambda) \in \mathbb{R}, \lambda = 1,2,\cdots,n.$ $\gamma(u_j^F, u^-) \Rightarrow \gamma(u_i^F, u^-)$ can be proved in a similar way as shown and it is omitted.

Corollary 1. *Let* $([u]_{F,i},[u]_i^F),([u]_{F,j},[u]_j^F)$ *be* F *-rough decision law generated from* $([u]_-,[u]^-)$ *respectively;* $(\alpha_-,\alpha^-),(\alpha_{F,i},\alpha_i^F)$, *and* $(\alpha_{F,j},\alpha_j^F)$ *be attribute sets of* ($[u]_-$, $[u]^-$), $([u]_{F,i},[u]_i^F)$, $([u]_{F,j},[u]_j^F)$ *respectively, if*

$$(\alpha_{F,j},\alpha_j^F) \Rightarrow (\alpha_{F,i},\alpha_i^F) \Rightarrow (\alpha_-,\alpha^-), \tag{29}$$

then
$$\gamma(u_{F,i},u_-) \le \gamma(u_{F,j},u_-)\,,\quad \gamma(u_i^F,u^-) \le \gamma(u_j^F,u^-)\quad . \tag{30}$$

Corollary 1 can be proved in a similar way as theorem 7 and it is omitted.

Theorem 8. *(The invariance theorem of* F *-rough decision law relation metric) Let* $([u]_{F,i},[u]_i^F)$ *and* $([u]_{F,j},[u]_j^F)$ *be* F *-rough decision law;* $(\alpha_{F,i},\alpha_i^F)$ *and* $(\alpha_{F,j},\alpha_j^F)$ *be the attribute sets of* $([u]_{F,i},[u]_i^F)$ *and* $([u]_{F,j},[u]_j^F)$ *respectively, and* $(\alpha_{F,j},\alpha_j^F) \ne (\alpha_{F,i},\alpha_i^F)$; *if* $\exists f, f' \in F, \bar{f}, \bar{f}' \in \bar{F}, \alpha_\lambda \in \alpha_{F,i}, \alpha'_\gamma \in \alpha_i^F, \bar{f}(\alpha_\lambda) = \beta_\lambda \,\bar{\in}\, \alpha_{F,i}, \bar{f}'(\alpha'_\gamma) = \beta_p \,\bar{\in}\, \alpha_{F,i}$, $\beta'_\gamma \,\bar{\in}\, \alpha_i^F, \beta'_q \,\bar{\in}\, \alpha_i^F, f(\beta_p) = \alpha_p \in \alpha_{F,i}$, $f(\beta'_q) = \alpha'_q \in \alpha_i^F$, *moreover*

$$\alpha_{F,j} = \alpha_{F,i} \cup \{\alpha_p \mid f(\beta_p) = \alpha_p\} - \{\alpha_\lambda \mid \bar{f}(\alpha_\lambda) = \beta_\lambda\} \tag{31}$$

$$\alpha_j^F = \alpha_i^F \cup \{\alpha'_q \mid f(\beta'_q) = \alpha'_q\} - \{\alpha_\gamma \mid \bar{f}'(\alpha_\gamma) = \beta_\gamma\} \tag{32}$$

then
$$\gamma(u_{F,i},u_-) = \gamma(u_{F,j},u_-)\,,\ \gamma(u_i^F,u^-) = \gamma(u_j^F,u^-)\,. \tag{33}$$

Proof. Since $\alpha_{F,i} \ne \alpha_{F,j}$, then $[u]_{F,i} \ne [u]_{F,j}$; Since $\exists f \in F, \bar{f} \in \bar{F}, \alpha_\lambda \in \alpha_{F,i}, \bar{f}(\alpha_\lambda) = \beta_\lambda \,\bar{\in}\, \alpha_{F,i}, \beta_p \,\bar{\in}\, \alpha_{F,i}, f(\beta_p) = \alpha_p \in \alpha_{F,i}$, moreover $\alpha_{F,j} = \alpha_{F,i} \cup \{\alpha_p \mid f(\beta_p) = \alpha_p\} - \{\alpha_\lambda \mid \bar{f}(\alpha_\lambda) = \beta_\lambda\}$, so $[u]_{F,i}$ and $[u]_{F,j}$ have the same attribute, $[u]_{F,i} = [u]_{F,j}$, namely $u_{F,i} = u_{F,j}$; by (18) -(22), it follows

$\gamma(u_{F,i}, u_-) = \gamma(u_{F,j}, u_-)$. $\gamma(u_i^F, u^-) = \gamma(u_j^F, u^-)$ can be proved in a similar way as shown (24) and it is omitted.

Theorem 9. *Let* $(\alpha_{F,i}, \alpha_i^F), (\alpha_{F,j}, \alpha_j^F)$ *and* $(\alpha_{F,k}, \alpha_k^F)$ *be attribute sets of* F *-rough decision law* $([u]_{F,i}, [u]_i^F), ([u]_{F,j}, [u]_j^F)$ *and* $([u]_{F,k}, [u]_k^F)$ *which are generated from* ($[u]_-, [u]^-$), *moreover*

$$(\alpha_{F,i}, \alpha_i^F) \Rightarrow (\alpha_{F,j}, \alpha_j^F) \Rightarrow (\alpha_{F,k}, \alpha_k^F) \tag{34}$$

Then $\mathrm{BAND}\{D_{F,t}, D_t^F\}$ composed of $([u]_{F,t}, [u]_t^F)$ generate

$$\mathrm{BAND}\{D_{F,k}, D_i^F\} \tag{35}$$

$\mathrm{BAND}\{D_{F,t}, D_t^F\}$ is sub-bands of $\mathrm{BAND}\{D_{F,k}, D_i^F\}$, Where $t = i, j, k$.

By theorems 7-9 and corollary 1, it follows:

The principle of F -rough decision law discovery

If there is $([u]_{F,k}, [u]_k^F)$ in $\{([u]_{F,i}, [u]_i^F) \mid i = 1, 2, \cdots, m\}$, and its F -rough decision law relation metric $(\gamma(u_{F,k}, u_-),\ \gamma(u_k^F, u^-))$ satisfies

$$\gamma(u_{F,\lambda}, u_-) = \max_{i=1}^{m}(\gamma(u_{F,i}, u_-)),\ \gamma(u_\lambda^F, u^-) = \max_{i=1}^{m}(\gamma(u_i^F, u^-)) \tag{36}$$

then F - rough decision law $([u]_{F,k}, [u]_k^F)$ is discovered in $\{([u]_{F,i}, [u]_i^F) \mid i = 1, 2, \cdots, m\}$.

5 The Application of F -Rough Decision Law

In order to facilitate discussion and without losing generality, let $[u]_- = \{u_1, u_2\}, [u]^- = \{u_1, u_2, u_3, u_4\}; [u]_{F,1} = \{u_1, u_5\}, [u]_1^F = \{u_1, u_3, u_4, u_5\}; [u]_{F,2} = \{u_2, u_6\}$, $[u]_2^F = \{u_2, u_3, u_4, u_6\}$; the origin- al data of $u_1 \sim u_6$ are omitted. The attribute set of $([u]_-, [u]^-)$ is $(\alpha_- = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}, \alpha^- = \{\alpha_1, \alpha_2, \alpha_3\})$, the attribute set of $([u]_{F,1}, [u]_1^F)$ is $(\alpha_{F,1} = \{\alpha_1, \alpha_2, \alpha_3, \alpha_5\}, \alpha_1^F = \{\alpha_1, \alpha_2, \alpha_5\})$ and the attribute set of $([u]_{F,2}, [u]_2^F)$ is $(\alpha_{F,2} = \{\alpha_1, \alpha_2, \alpha_4, \alpha_6\}, \alpha_2^F = \{\alpha_1, \alpha_2, \alpha_6\})$; the names of attribute sets $\alpha_1 \sim \alpha_6$ are omitted. $u_{F,1}, u_{F,2}$ and u_- are the composition of $[u]_{F,1}, [u]_{F,2}$ and $[u]_-$ respectively; u_1^F , u_2^F and u^- are respectively the composition of $[u]_1^F$, $[u]_2^F$ and $[u]^-$. The standardized values of $u_{F,1}, u_{F,2}$ and u_- can be found in table 1. The original data (The data of *k*>6 are omitted) $u_{F,1}$, $u_{F,2}$ and u_- which have been standardized won't affect the analysis of approach given by the example. The example l in this section is from the profit risk analysis of investment system.

By (18)-(22), table 2 and table 3 can be obtained. ρ is the detached coefficient and $\rho = 1$; by (19), table 5 can be obtained, since $\max_{i,k} \Delta_{0,i}(k) = \Delta_{0,2}(3) = 1.7810$. By shown in table 4. By (22), we can obtain: $\gamma(u_{F,1}, u_{\_}) = 0.7140$, $\gamma(u_{F,2}, u_{\_}) = 0.6562$. Similarly to the former computational process, by the standardized data of u_1^F, u_2^F, u^- and (18)-(22), it follows: $\gamma(u_1^F, u^-) = 0.6735$, $\gamma(u_2^F, u^-) = 0.5641$.

The standardized data and computational process of u_1^F, u_2^F and u^- are omitted. We can obtain the F -rough decision law relation metric $(\gamma(u_{F,1}, u_{\_}), \gamma(u_1^F, u^-)) = (0.7140, \quad 0.6735), (\gamma(u_{F,2}, u_{\_}), \gamma(u_2^F, u^-)) = 0.5641)$ respectively, so that $\gamma(u_{F,2}, u_{\_}) \le \gamma(u_{F,1}, u_{\_}), \gamma(u_2^F, \quad u^-) \le \gamma(u_1^F, u^-), \gamma(u_2^F, u^-) \le \gamma(u_1^F, u^-)$, namely $(\gamma(u_{F,2}, u_{\_}), \gamma(u_2^F, u^-)) \le (\gamma(u_{F,1}, u_{\_}), \gamma(u_1^F, u^-))$. According to the discussion above, We can conclude that F -rough decision law $([u]_{F,1}, \ [u]_1^F)$ is closer to rough decision law $([u]_{\_}, [u]^-)$ than F -rough decision law $([u]_{F,2}, [u]_2^F)$. If the profit decision law which is denoted by F -rough decision law $([u]_{F,1}, [u]_1^F)$ has little profit risk. This conclusion is affirmed in the profit risk decision analysis of real investment system; F -rough decision law $([u]_{F,1}, [u]_1^F)$ can be discovered in $\{([u]_{F,1}, [u]_1^F), \ \ ([u]_{F,2}, [u]_2^F)\}$.

Table 1. The Discrete Standardized Data of F -decision Law $u_{F,1}, u_{F,2}$ and $u_{\_}$.

k	1	2	3	4	5	6	…
$u_{\_}$	1.0000	1.5176	1.7602	2.4567	3.6105	4.7536	…
$u_{F,1}$	1.0000	1.7642	1.6946	2.9034	3.2642	5.4370	…
$u_{F,2}$	1.0000	1.9126	1.8253	2.0648	3.8553	4.2916	…

Table 2. The Values Distribution of $P_{0,i}(k), S_{0,i}(k)$, $i = 1, 2; k = 1 \sim 6, \cdots$.

k	*1*	*2*	*3*	*4*	*5*	*6*	…
$P_{0,1}(k)$	0	0.2466	0.0656	0.4467	0.3463	0.6834	…
$P_{0,2}(k)$	0	0.3950	0.0651	0.3919	0.2448	0.4620	…
$S_{0,1}(k)$	0.2466	0.3122	0.5123	0.7930	1.0297	—	…
$S_{0,2}(k)$	0.3950	0.3299	0.4570	0.6367	0.7068	—	…

"—"in table 2 means that there is no data.

Table 3. The Value Distribution of $\overline{S}_{0,i}(k), i=1,2; k=1 \sim 6,\cdots$.

k	1	2	3	4	5	6	…
$\overline{S}_{0,1}(k)$	0.1233	0.1043	0.1990	0.2014	0.2850	—	…
$\overline{S}_{0,2}(k)$	0. 1975	0.2300	0.1727	0.1677	0.1934	—	…

Table 4. The Value Distribution of $\Delta_{0,i}(k), i=1,2; k=1 \sim 6,\cdots$.

k	1	2	3	4	5	6	…
$\Delta_{0,1}(k)$	0.3699	0.4165	0.7113	0.9944	0.6983	—	…
$\Delta_{0,2}(k)$	0.5925	0.5600	0.6297	0.8044	0.9002	—	…

Table 5. The Value Distribution of $\mu(u_{F,i}(k), u\_(k))$, $i=1,2; k=1 \sim 6,\cdots$

Where $\mu_1 = \mu(u_{F,1}(k), u\_(k)), \mu_2 = \mu(u_{F,2}(k), u\_(k))$.

k	1	2	3	4	5	6	…
μ_1	0.8280	0.8105	0.7146	0.6417	0.5753	—	…
μ_2	0.6893	0.7013	0.6762	0.6204	0.5936	—	…

6 Conclusions

Which has dynamic characteristic and decision law characteristic is an important approach to seek (or mine) the unknown decision law in system, and it has an application of a good future. By using function two direction S-rough sets, this paper presents the discussion of F -rough decision law and application. Rough decision law comes from such a background: in economic decision-making system, profit decision analysis curve (decision analysis law) shouldn't be indicated by just one, but by two curves (they compose the rough decision law), since the curve is fluctuating for it is always attacked by the attribute (the variation of the investment situation) and then it forms two curves which are called the minimum profit decision analysis law and the top gain decision analysis law; this fact is just included in the lower approximation $(R,F)_{\circ}(Q^*)$ and the upper approximation $(R,F)^{\circ}(Q^*)$ of function two direction S-rough sets. Function S-rough sets develops the research of Z. Pawlak rough sets theory [11] and its application.

Acknowledgements. The work was supported by Fujian Provincial Key Laboratory of imitation Brain intelligent system Open topics (BL I SSOS 2010104), the Science and Technology Foundation of Education Department of Fujian Province, China (JA09232, JA08234); and Science and Technology Develop Item of Sanming University (HX201007/Z).

References

1. Shi, K.: Function S-rough sets and function transfer. An International Journal Advances in Systems Science and Applications 1, 1–8 (2005)
2. Shi, K.: S-rough sets and its applications in diagnosis-recognition for disease. In: IEEE Proceedings of the First International Conference on Machine Learning and Cybernetics, vol. 1, pp. 50–54 (2002)
3. Shi, K., Cu, Y.: S-rough sets and rough decision, pp. 53–56. Science Press, Beijing (2006)
4. Shi, K., Yao, B.: Function S-rough set and system decision law mining, pp. 180–185. Science Press, Beijing (2007)
5. Zhang, P., Shi, K.: Function S-rough sets and rough decision law heredity-mining. In: IEEE proceedings of the Fourth International Conference on Machine Learning and Cybernetics, vol. 3, pp. 3148–3152 (2005)
6. Shi, K., Yao, B.: Function S-rough sets and recognition of financial risk decision laws. In: The First International Conference on Rough Sets and Knowledge Technology, vol. 1, pp. 247–253 (2006)
7. Shi, K., Xia, J.: Function S-rough sets and mining-discovery of rough decision laws in system. Journal of Systems Engineering and Electronics 4, 919–926 (2006)
8. Shi, K., Huang, Y.-p.: Function S-rough sets and decision law mining. In: International Conference on Fuzzy Information Engineering, vol. 1, pp. 48–57 (2007)
9. Shi, K., Yao, B.: Function S-rough sets and decision law identification. Science in China Series F: Information Sciences 51, 499–510 (2008)
10. Shi, K., jianli, Z.: Function S-rough sets and security-authentication of hiding decision law. Science in China Series F: Information Sciences 51, 924–935 (2008)
11. Pawlak, Z.: Rough sets. International Journal of Computer and Information Sciences 11, 341–356 (1982)

On ⊙-Ideals and Lattices of ⊙-Ideals in Regular Residuated Lattices

Chun-hui Liu[1,2] and Luo-shan Xu[1,⋆]

[1] Department of Mathematics, Yangzhou University, Yangzhou 225002, China
e-mail: luoshanxu@hotmail.com

[2] Department of Elementary Education, Chifeng University, Chifeng 024001, China
e-mail: chunhuiliu1982@163.com

Abstract. In this paper, the operation ⊙ and the concept of ⊙-ideals of (regular) residuated lattices are introduced. Some characterization theorems for ⊙-ideals of (regular) residuated lattices are given. Representation theorems about ⊙-ideals which are generated by non-empty subsets of regular residuated lattices are obtained. For the set of all ⊙-ideals of a (regular) residuated lattice, an adjunction pair is defined. It is proved that the lattice of all ⊙-ideals in a regular residuated lattice with the adjunction and the set-inclusion order is a complete Heyting algebra (i.e., a frame) and an algebraic lattice, which thus gives a new distributive residuated lattice.

Keywords: Residuated lattice; ⊙-ideal; algebraic lattice; frame.

1 Introduction

With the developments of mathematics and computer science, non-classical mathematical logic has become a formal tool for artificial intelligence to deal with uncertainty information. One of important branch of non-classical mathematical logic is to study logic algebra systems. Results in this area not only promoted the development of non-classical mathematical logic, but also enriched the contents of algebra [1, 2]. Among various logic algebra systems, residuated lattices introduced by Pavelka are important ones and are reasonable extensions of Heyting algebras. Based on the Łukasiewicz axiom system, Pavelka introduced the theory of residuated lattices into the studies of non-classical mathematical logic, established a kind of logic construction, and solved the completeness problem of Łukasiewicz axiom system[3]. Residuated lattices have been considered a kind of idealistic algebras in the theory of non-classical mathematical logic. It is worthy of noting that various logic algebras based on different implication operators, such as MV-algebras,

⋆ Corresponding author.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 425–434.
springerlink.com

BL-algebras, R_0-algebras and lattice implication algebras, etc., are special residuated lattices[4, 5]. Thus, it is meaningful to deeply study properties of residuated lattices.

Ideals play important roles in studying various reasoning systems and logic algebras. Properties of types of ideals in logic algebras have been actively and deeply studied[6]$^-$[11]. In this paper, the concepts of the operation $\odot$ and $\odot$-ideals of residuated lattices are introduced and properties and characterizations of them are discussed. For the set of all $\odot$-ideals of a (regular) residuated lattice, an adjunction pair is defined. It is proved that the lattice of all $\odot$-ideals on a regular residuated lattice with the adjunction and the set-inclusion order is an algebraic lattice and a complete Heyting algebra (i.e., a frame). A new distributive residuated lattice is thus induced.

2 Basic Notions and Related Results

We in this section recall some basic notions and related results needed in the sequel. The non-explicitly stated notions on posets and domain theory, please refer to [12].

Definition 2.1 [1]**.** *Let P be a poset. The two binary operations $\otimes$ and $\rightarrow$ on P are said to be adjoint to each other, if they satisfy the following conditions* (1)–(3):

(1) $\otimes : P \times P \rightarrow P$ *is isotone*;

(2) $\rightarrow: P \times P \rightarrow P$ *is antitone in the first variable and isotone in the second variable;*

(3) $a \otimes b \leq c$ *if and only if* $a \leq b \rightarrow c$, *for all* $a, b, c \in P$.

And $(\otimes, \rightarrow)$ *is called an adjoint pair on* P.

Definition 2.2 [3, 13]**.** *A structure $L = (L; \leq, \otimes, \rightarrow, 0, 1)$ is called a residuated lattice if the following conditions satisfied.*

(1) $(L, \leq)$ is a bounded lattice, 0 is the smallest element and 1 is the greatest element of L, respectively;

(2) $(L, \otimes, 1)$ is a commutative semigroup with unit element 1;

(3) $(\otimes, \rightarrow)$ is an adjoint pair on L.

Lemma 2.1 [13]**.** *Let L be a residuated lattice. Then for all $a, b, c \in L$ we have*

(R1) $a \leq b \rightarrow a \otimes b$ *or* $a \wedge (b \rightarrow a \otimes b) = a$;

(R2) $a \leq b$ *if and only if* $a \rightarrow b = 1$;

(R3) $(a \rightarrow b) \otimes a \leq b$ *or* $((a \rightarrow b) \otimes a) \vee b = b$;

(R4) $1 \rightarrow a = a$;

(R5) $(a \vee b) \otimes c = (a \otimes c) \vee (b \otimes c)$;

(R6) $a \otimes b \leq a \wedge b$;

(R7) $b \rightarrow c \leq (a \rightarrow b) \rightarrow (a \rightarrow c)$;

(R8) $a \leq b \rightarrow c$ *if and only if* $b \leq a \rightarrow c$;

(R9) $a \rightarrow (b \rightarrow c) = b \rightarrow (a \rightarrow c)$;

(R10) $a \to b \leq a \otimes c \to b \otimes c$;
(R11) $a \otimes b \to c = a \to (b \to c)$;
(R12) $a \to b \leq (b \to c) \to (a \to c)$;
(R13) $a \to (b \to a) = 1$;
(R14) $a \to b \leq a \wedge c \to b \wedge c$;
(R15) $a \to b \leq a \vee c \to b \vee c$.

Definition 2.3 [13]**.** *Let L be a residuated lattice. Define on L the unary operation $\neg : L \to L$ such that $\neg a = a \to 0$ for any $a \in L$. Then $\neg$ is called the pseudo-complement operator of L. And we call L a regular residuated lattice if $\neg\neg a = a$ for any $a \in L$.*

Remark 2.1. *For a residuated lattice L and $\forall a, b, c \in L$, it is easy to show that*
(1) $\neg 0 = 1, \quad \neg 1 = 0$;
(2) $a \leq \neg a$ *and* $\neg\neg\neg a = \neg a$;
(3) $a \leq b$ *implies* $\neg b \leq \neg a$.

Lemma 2.2 [13]**.** *For a regular residuated lattice L and $\forall a, b, c \in L$ we have*
(RR1) $\neg a \to \neg b = b \to a$ *and* $a \to \neg b = b \to \neg a$ *and* $\neg a \to b = \neg b \to a$;
(RR2) $a \otimes b = \neg(a \to \neg b)$ *and* $a \to b = \neg(a \otimes \neg b)$;
(RR3) $a \otimes \neg a = 0$;
(RR4) $\neg a \to (a \to b) = 1$.

3 The ⊙-Operation and ⊙-Ideals in Residuated Lattices

We in this section introduce the operations $\oplus$ and $\odot$ on a residuated lattice L and discuss their properties. Then define $\odot$-ideals of a residuated lattice L.

Definition 3.1. *Let L be a residuated lattice. Binary operations $\oplus$ and $\odot$ on L are defined as following: $a \oplus b = \neg a \to b$ and $a \odot b = \neg(a \to b)$, for all $a, b \in L$.*

Lemma 3.1. *Let L be a regular residuated lattice. Then for all $a, b, c \in L$ we have*
(RR5) $0 \odot a = 0$ *and* $a \odot 0 = a$;
(RR6) $(a \odot b) \odot c = (a \odot c) \odot b$;
(RR7) $a \leq b$ *if and only if* $a \odot b = 0$, *in particular,* $a \odot a = 0$;
(RR8) *If* $a \leq b$, *then* $a \odot c \leq b \odot c$ *and* $c \odot b \leq c \odot a$;
(RR9) $a \odot (a \odot (a \odot b)) = a \odot b$;
(RR10) $(a \wedge c) \odot (b \wedge c) \leq a \odot b$;
(RR11) $a \oplus b = b \oplus a$ *and* $(a \oplus b) \oplus c = a \oplus (b \oplus c)$;
(RR12) $(c \oplus a) \odot (c \oplus b) \leq a \odot b$;
(RR13) $a \leq b$ *implies* $a \oplus c \leq b \oplus c$;

(RR14) $(a \wedge b) \oplus c = (a \oplus c) \wedge (b \oplus c)$ *and* $(a \vee b) \odot c = (a \odot c) \vee (b \odot c)$*;*
(RR15) $c \odot (a \oplus b) = (c \odot b) \odot a$*;*
(RR16) $a \wedge (b \oplus c) \leq (a \wedge b) \oplus (a \wedge c)$*;*
(RR17) $a \oplus b \geq a \vee b$*.*

Proof. (RR5): $0 \odot a = \neg(0 \to a) = \neg 1 = 0$ and $a \odot 0 = \neg(a \to 0) = \neg\neg a = a$.

(RR6): By (RR1), (RR9) and the regular property of L we have that

$$\begin{aligned}(a \odot b) \odot c &= \neg(\neg(a \to b) \to c) = \neg(\neg c \to \neg\neg(a \to b))\\ &= \neg(\neg c \to (\neg b \to \neg a))\\ &= \neg(\neg b \to (\neg c \to \neg a))\\ &= \neg(\neg(a \to c) \to b) = (a \odot c) \odot b.\end{aligned}$$

(RR7): By (R8) we have $a \leq b \Leftrightarrow a \to b = 1 \Leftrightarrow \neg(a \to b) = \neg 1 = 0 \Leftrightarrow a \odot b = 0$.

(RR8): Assume $a \leq b$, then by Definition 2.1(2) we have that $b \to c \leq a \to c$, thus $(b \to c) \to (a \to c) = 1$. Combine this with (RR1) we obtain that $(a \odot c) \to (b \odot c) = \neg(a \to c) \to \neg(b \to c) = (b \to c) \to (a \to c) = 1$. Using (R8) again we have that $a \odot c \leq b \odot c$. Similarly, we can obtain that $c \odot b \leq c \odot a$.

(RR9): On one hand, since $(a \odot (a \odot (a \odot b))) \odot (a \odot b) = (a \odot (a \odot b)) \odot (a \odot (a \odot b)) = 0$, by (RR6) and (RR7), we have that $a \odot (a \odot (a \odot b)) \leq a \odot b$. On the other hand, since $(a \odot (a \odot b)) \odot b = (a \odot b) \odot (a \odot b) = 0$, by (RR7) we have $a \odot (a \odot b) \leq b$, and $a \odot (a \odot (a \odot b)) \geq a \odot b$ by (RR8). To sum up the above two hands, we know (RR9) holds.

(RR10): By (R14) and Remark 2.1(3) we have that $(a \wedge c) \odot (b \wedge c) = \neg(a \wedge c \to b \wedge c) \leq \neg(a \to b) = a \odot b$.

(RR11): It is immediately follows from the definition of $\oplus$ and (RR1).

(RR12): It follows from (R7) that $(c \oplus a) \odot (c \oplus b) = \neg((\neg c \to a) \to (\neg c \to b)) \leq \neg(a \to b) = a \odot b$.

(RR13): Assume $a \leq b$. Then $a \odot b = 0$ by (RR7). Since $(a \oplus c) \odot (b \oplus c) \leq a \odot b$ by (RR11) and (RR12), we have that $(a \oplus c) \odot (b \oplus c) = 0$, hence $a \oplus c \leq b \oplus c$.

(RR14): By (R5), (RR1), (RR2) and the regular property of L we obtain that

$$\begin{aligned}(a \wedge b) \oplus c &= \neg(a \wedge b) \to \neg\neg c = \neg(\neg(a \wedge b) \otimes \neg c)\\ &= \neg((\neg a \vee \neg b) \otimes \neg c) = \neg((\neg a \otimes \neg c) \vee (\neg b \otimes \neg c))\\ &= \neg(\neg a \otimes \neg c) \wedge \neg(\neg b \otimes \neg c) = \neg\neg(\neg a \to c) \wedge \neg\neg(\neg b \to c)\\ &= (a \oplus c) \wedge (b \oplus c).\end{aligned}$$

By (RR2) we have that $x \odot y = x \otimes \neg y$ for all $x, y \in L$. It follows from (R5) that

$$(a \vee b) \odot c = (a \vee b) \otimes \neg c = (a \otimes \neg c) \vee (b \otimes \neg c) = (a \odot c) \vee (b \odot c).$$

So, the two equations in (RR14) hold.

(RR15): By (R9) and (RR1) we have that

$$\begin{aligned}c \odot (a \oplus b) &= \neg(c \to (\neg a \to b)) = \neg(\neg a \to (c \to b))\\ &= \neg(\neg(c \to b) \to a) = \neg((c \odot b) \to a) = (c \odot b) \odot a.\end{aligned}$$

(RR16): Since by (RR5), (RR6), (RR7) and (RR15) that

$$\begin{aligned} &a \wedge (b \oplus c)) \odot (a \oplus (b \wedge c)) = ((a \wedge (b \oplus c)) \odot (b \wedge c)) \odot a[\text{by (RR15)}] \\ &= ((a \wedge (b \oplus c)) \odot a) \odot (b \wedge c)[\text{by (RR6)}] \\ &= 0 \odot (b \wedge c)[\text{by (RR7)}] \\ &= 0, [\text{by (RR5)}] \end{aligned}$$

we can obtain that $a \wedge (b \oplus c) \leq a \oplus (b \wedge c)$. Combine this with (RR14) we have that

$$\begin{aligned} &(a \wedge b) \oplus (a \wedge c) = (b \oplus c) \wedge (a \oplus c) \wedge (a \oplus b) \wedge (a \oplus a) \\ &= (a \oplus (b \wedge c)) \wedge (b \oplus c) \wedge (a \oplus a) \\ &\geq (a \wedge (b \oplus c)) \wedge (b \oplus c) \wedge a \\ &= a \wedge (b \oplus c). \end{aligned}$$

So, (RR16) holds.

(RR17): By Definition 2.1(2) we have that $a \oplus b = \neg a \to b \geq \neg a \to 0 = \neg\neg a = a$. And by (RR11) we have that $a \oplus b \geq b$. So, $a \oplus b \geq a \vee b$.

Definition 3.2. *Let L be a residuated lattice. A nonempty subset I of L is called a $\odot$-ideal of L if it satisfies the following conditions:*

(I1) $0 \in I$;

(I2) *If $y \in I$ and $x \odot y \in I$, then $x \in I$, for all $x, y \in L$.*

The set of all $\odot$-ideals of L is denoted by $\odot\mathrm{I}(L)$.

Remark 3.1. *(1) Let L be a residuated lattice. It is easy to checked that $\{0\}, L \in \odot I(L)$. If $\mathcal{I}$ is a non-empty family of $\odot$-ideals of L, Then $\cap\mathcal{I}$ is also a $\odot$-ideal of L.*

(2) Let L be a regular residuated lattice, $I \in \odot I(L)$ and $a, b \in L$. If $a \leq b$ and $b \in I$, then $a \in I$, i.e., any $\odot$-ideals of L is a down set. In fact, if $a \leq b$, then $a \odot b = 0 \in I$ by (RR5), hence $a \in I$ by $b \in I$.

(3) Let L be a regular residuated lattice and $I \in \odot I(L)$. Then I is right $\odot$-closed with elements of L. In fact, assume that $a \in I$, $x \in L$. Since $(a \odot x) \odot a = (a \odot a) \odot x = 0 \odot x = 0 \in I$, by (RR6), it follows from $I \in \odot I(L)$ and $a \in I$ that $a \odot x \in I$.

(4) Let L be a regular residuated lattice and $I \in \odot I(L)$. Then I is closed with binary $\vee$. In fact, assume $a, b \in I$, then by (RR14) and (3) we have that $(a \vee b) \odot a = (a \odot a) \vee (b \odot a) = b \odot a \in I$. So, $a \vee b \in I$ by $I \in \odot I(L)$.

Theorem 3.1. *Let L be a regular residuated lattice and I a nonempty subset of L. Then $I \in \odot I(L)$ if and only if I is a down set and closed with operation $\oplus$.*

Proof. $\Leftarrow$ Since I is a down set, we know that $0 \in I$. Assume $a \in L, b \in I$ and $a \odot b \in I$, then it follows from the closedness of I with operation $\oplus$ that $b \oplus (a \odot b) \in I$. By (RR1), (R2), (R9) and the regular property of L we have that

$$a \to (b \oplus (a \odot b)) = a \to (\neg b \to \neg(a \to b)) = a \to ((a \to b) \to b) = (a \to b) \to (a \to b) = 1.$$

This shows that $a \leq b \oplus (a \odot b)$. Since I is a down set and $b \oplus (a \odot b) \in I$, we obtain that $a \in I$ and $I \in \odot\mathrm{I}(L)$.

$\Rightarrow$ Assume $I \in \odot\mathrm{I}(L)$. Then I is a down set by Remark 3.1(2). Now we show that I is closed with operation $\oplus$. In fact, for all $a, b \in I$, by (RR1), (R9) and the regular property of L we can obtain that

$$\begin{aligned}((a \oplus b) \odot b) \odot a &= \neg(\neg((a \oplus b) \to b) \to a) \\ &= \neg(\neg a \to ((a \oplus b) \to b)) = \neg((a \oplus b) \to (\neg a \to b)) \\ &= \neg((a \oplus b) \to (a \oplus b)) = \neg 1 = 0 \in I \in \odot\mathrm{I}(L).\end{aligned}$$

So, $a \oplus b \in I$ for $a, b \in I$ by Definition 3.2. So, I is closed with operation $\oplus$.

The following Corollary is immediate by Theorem 3.1 and (RR17). But the converse may not be true and Example 4.1 serves as a counterexample.

Corollary 3.1. *Every $\odot$-ideal of a regular residuated lattice is an ordinary ideal.*

4 Generated $\odot$-Ideals by a Nonempty Subset

Definition 4.1. *Let L be a regular residuated lattice and A a nonempty subset of L. The least $\odot$-ideal containing A is called the $\odot$-ideal generated by A, denoted by $\langle A \rangle$.*

It is obvious that $\langle A \rangle = \bigcap_{I \in \odot\mathrm{I}(L), A \subset I} I$ and $A \subset \langle A \rangle$. If $I \in \odot\mathrm{I}(L)$, then $\langle I \rangle = I$.

Theorem 4.1. *If A is a nonempty subset of a regular residuated lattice L, then*

$$\begin{aligned}\langle A \rangle &= \{x \in L | (\cdots((x \odot a_1) \odot a_2) \odot \cdots) \odot a_n = 0,\ \exists\, a_1, a_2, \cdots, a_n \in A \text{ and } n \in \mathbf{N}\} \\ &= \{x \in L | x \le a_1 \oplus a_2 \oplus \cdots \oplus a_n,\ \exists\, a_1, a_2, \cdots, a_n \in A \text{ and } n \in \mathbf{N}\}.\end{aligned}$$

Proof. For the sake of convenience, let

$$B = \{x \in L | (\cdots((x \odot a_1) \odot a_2) \odot \cdots) \odot a_n = 0,\ \exists\, a_1, a_2, \cdots, a_n \in A \text{ and } n \in \mathbf{N}\}$$

and

$$C = \{x \in L | x \le a_1 \oplus a_2 \oplus \cdots \oplus a_n,\ \exists\, a_1, a_2, \cdots, a_n \in A \text{ and } n \in \mathbf{N}\}.$$

Firstly, we show that $B = \langle A \rangle$. Obviously, $A \subset B$ and $0 \in B$. Now assume $y \in B$ and $x \odot y \in B$. Then there exist $a_1, a_2, \cdots, a_n \in A$ and $b_1, b_2, \cdots, b_m \in A$ such that

$$(\cdots((y \odot a_1) \odot a_2) \odot \cdots) \odot a_n = 0 \text{ and } (\cdots(((x \odot y) \odot b_1) \odot b_2) \odot \cdots) \odot b_m = 0.$$

Thus by (RR6) and (RR7) we have that $(\cdots((x \odot b_1) \odot b_2) \odot \cdots) \odot b_m \le y$, hence

$$\begin{aligned}0 &= (\cdots((y \odot a_1) \odot a_2) \odot \cdots) \odot a_n \\ &\ge (\cdots((((\cdots((x \odot b_1) \odot b_2) \odot \cdots) \odot b_m) \odot a_1) \odot a_2) \odot \cdots) \odot a_n\end{aligned}$$

by (RR8). Thus $x \in B$ by the definition of B. So, $B \in \odot\mathrm{I}(L)$ and $\langle A \rangle \subset B$.

Since for any $z \in B$, there exists some $c_1, c_2, \cdots, c_k \in A \subset \langle A \rangle$ such that

$$(\cdots((z \odot c_1) \odot c_2) \odot \cdots) \odot c_k = 0,$$

we have that $z \in \langle A \rangle$ by $\langle A \rangle \in \odot\mathrm{I}(L)$. Thus $B \subset \langle A \rangle$. Therefore $B = \langle A \rangle$.

Secondly, we show that $B = C$. By (RR7) we have that $x \leq a_1 \oplus a_2 \oplus \cdots \oplus a_n$ if and only if $x \odot (a_1 \oplus a_2 \oplus \cdots \oplus a_n) = 0$, and by (RR8) we have that $x \odot (a_1 \oplus a_2 \oplus \cdots \oplus a_n) = (x \odot a_1) \odot (a_2 \oplus \cdots \oplus a_n) = \cdots = (\cdots((x \odot a_1) \odot a_2) \odot \cdots) \odot a_n$. Thus, by the definitions of the sets B and C we have that $B = C$.

Theorem 4.2. *Let L be a regular residuated lattice and $\{A_i\}$ a directed family of subsets of L. Then $\langle \bigcup A_i \rangle = \bigcup \langle A_i \rangle$. Therefore the union of any directed family of ⊙-ideals of L is also a ⊙-ideal.*

Proof. It is sufficient to show that $\bigcup \langle A_i \rangle$ is a ⊙-ideals of L. Obviously, by the directed property of $\{A_i\}$ and Theorem 3.1 we have $\bigcup \langle A_i \rangle \in \odot\mathrm{I}(L)$.

Example 4.1. *Let $L = \{0, a, b, c, d, 1\}$ and Hasse diagram of L is given as Fig. 4.1. Operators $\rightarrow$ and $\otimes$ of L are defined in Table 4.1 and Table 4.2, respectively. Operators $\neg, \odot$ and $\oplus$ are defined in Table 4.3, Table 4.4and Table 4.5, respectively as following:*

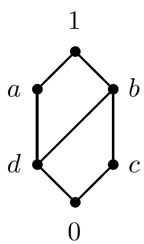


Fig. 4.1 Hasse Diag. of L

Table 4.1. Def. of $\rightarrow$

$\rightarrow$	0	a	b	c	d	1
0	1	1	1	1	1	1
a	c	1	b	c	b	1
b	d	a	1	b	a	1
c	a	a	1	1	a	1
d	b	1	1	b	1	1
1	0	a	b	c	d	1

Table 4.2. Def. of $\rightarrow$

$\otimes$	0	a	b	c	d	1
0	0	0	0	0	0	0
a	0	a	d	0	d	a
b	0	d	c	c	0	b
c	0	0	c	c	0	c
d	0	d	0	0	0	d
1	0	a	b	c	d	1

Table 4.3. Def. of $\neg$

x	$\neg x$
0	1
a	c
b	d
c	a
d	b
1	0

Table 4.4. Def. of $\odot$

$\odot$	0	a	b	c	d	1
0	0	0	0	0	0	0
a	a	0	d	a	d	0
b	b	c	0	d	c	0
c	c	c	0	0	c	0
d	d	0	0	d	0	0
1	1	c	d	a	b	0

Table 4.5 Def. of $\oplus$

$\oplus$	0	a	b	c	d	1
0	0	a	b	c	d	1
a	a	a	1	1	a	1
b	b	1	1	b	1	1
c	c	1	b	c	b	1
d	d	a	1	b	a	1
1	1	1	1	1	1	1

Then $(L, \rightarrow, \otimes)$is a regular residuated lattice. It is easy to check that $I = \{0, c\} \in \odot I(L)$. Since $A = \{0, a\}$ is not a down set, we know that $A \notin \odot I(L)$ and $\langle A \rangle = \{0, a, d\}$ is the generated ⊙-ideal by A. It is also easy to know that the principal ideal $\downarrow b$ of L is not a ⊙-ideal, for $b \oplus d = 1 \notin \downarrow b$. As a matter of fact, L has only four ⊙-ideals.

5 The ⊙-Ideal Lattice of a Regular Residuated Lattice

Theorem 5.1. *Let L be a regular residuated lattice. Then $(\odot I(L), \subset)$ is a distributive complete lattice. And directed sups are set-unions.*

Proof. Obviously, $(\odot\mathrm{I}(L), \subset)$ is a poset, $\{0\}$ is the smallest element and L is the greatest element of $\odot\mathrm{I}(L)$, respectively. By remark 3.1(1) we know that the infimum of any family of ⊙-ideals is just the set-join of the family. So, any subset of poset $(\odot\mathrm{I}(L), \subset)$ has a infimum, hence $(\odot\mathrm{I}(L), \subset)$ is a complete lattice. In particular, for all $A, B \in \odot\mathrm{I}(L)$, it is easy to show that $A \wedge B = A \cap B$ is the infimum of A and B, and $A \vee B = \langle A \cup B\rangle$ is the supremum of A and B. By Theorem 4.2 we know that the union of any directed family of ⊙-ideals of L is also a ⊙-ideal which is just the supremum of that family. That is to say, directed sups in $(\odot\mathrm{I}(L), \subset)$ are just set-unions.

Now we show that distributive laws hold. Since $(\odot\mathrm{I}(L), \subset)$ is a lattice, it is sufficient to show that $\forall A, B, C \in \odot\mathrm{I}(L)$, $A \wedge (B \vee C) = (A \wedge B) \vee (A \wedge C)$. Obviously, $A \wedge (B \vee C) \supset (A \wedge B) \vee (A \wedge C)$ and we only need to show that $A \wedge (B \vee C) \subset (A \wedge B) \vee (A \wedge C)$. In fact, assume $x \in A \wedge (B \vee C)$. Then $x \in A$ and $x \in B \vee C$. There exists some $z_1, z_2, \cdots, z_n \in B \cup C$ such that $x \le z_1 \oplus z_2 \oplus \cdots \oplus z_n$ by Theorem 4.1. And by (RR16) we have that

$$x = x \wedge (z_1 \oplus z_2 \oplus \cdots \oplus z_n) \le (x \wedge z_1) \oplus (x \wedge z_2) \oplus \cdots \oplus (x \wedge z_n).$$

Since A, B, C are down sets, we can obtain that

$$x \wedge z_i \in (A \cap B) \cup (A \cap C) = (A \wedge B) \cup (A \wedge C), \quad i = 1, 2, \cdots, n.$$

So by Theorem 4.1 we have $x \in (A \wedge B) \vee (A \wedge C)$ and $A \wedge (B \vee C) \subset (A \wedge B) \vee (A \wedge C)$. Thus $A \wedge (B \vee C) = (A \wedge B) \vee (A \wedge C)$, finishing the proof.

Theorem 5.2. *Let L be a regular residuated lattice and $A, B \in \odot I(L)$. Define $A \otimes B = A \wedge B$ and $A \to B = \{x \in L |\ a \wedge x \in B$ for all $a \in A\}$. Then $(\odot I(L), \otimes, \to)$ is a residuated lattice.*

Proof. Obviously, $A \otimes B \in \odot\mathrm{I}(L)$ and $(\odot\mathrm{I}(L), \otimes, L)$ is a commutative semi-group with unital element L. By Theorem 5.1 we know that $(\odot\mathrm{I}(L), \subset)$ is a bounded lattice. In order to prove that $(\odot\mathrm{I}(L), \otimes, \to)$ is a residuated lattice, it is sufficient to show that $(\otimes, \to)$ is an adjoint pair on $\odot\mathrm{I}(L)$ in the sense of Definition 2.1. At first, we prove that $A \to B \in \odot\mathrm{I}(L)$. In fact, since $\forall a \in A, a \wedge 0 = 0 \in B$, we have that $0 \in A \to B$. Assume $y \in A \to B$ and $x \le y$. Then $a \wedge y \in B$ and $a \wedge x \le a \wedge y$ for all $a \in A$. Thus by that $B \in \odot\mathrm{I}(L)$ is a down set we can obtain that $a \wedge x \in B$ and $x \in A \to B$. This means that $A \to B$ is a down set. Suppose that $x, y \in A \to B$, then $a \wedge x \in B$ and $a \wedge y \in B$ for any $a \in A$. It follows from $B \in \odot\mathrm{I}(L)$ and Theorem 3.1 that $(a \wedge x) \oplus (a \wedge y) \in B$. Thus by (RR16) and B is a down set we have $a \wedge (x \oplus y) \in B$ and $x \oplus y \in A \to B$. This means that $A \to B$ is closed with $\oplus$. Therefore $A \to B \in \odot\mathrm{I}(L)$ by Theorem 3.1.

It is obvious that $\otimes = \wedge$ is isotone and $A \subset B$ implies $B \to C \subset A \to C$ and $C \to A \subset C \to B$ for all $A, B, C \in \odot\mathrm{I}(L)$. Now we prove that $\forall A, B, C \in$

$\odot\mathrm{I}(L)$, $A \otimes B \subset C$ if and only if $A \subset B \to C$. In fact, suppose that $A \otimes B \subset C$ and $x \in A$, then $b \wedge x \leq x$ and $b \wedge x \leq b$ for any $b \in B$. Since A and B are down sets we have that $b \wedge x \in A \wedge B = A \otimes B \subset C$, $x \in B \to C$, and $A \subset B \to C$. Conversely, suppose $A \subset B \to C$ and $x \in A \otimes B = A \wedge B$. Then $x \in B \to C$ and $b \wedge x \in C$ for any $b \in B$ Put $b = x$ we have that $x = x \wedge x \in C$ and $A \otimes B \subset C$. This means that $(\otimes, \to)$ is a adjoint pair on $\odot\mathrm{I}(L)$. Therefore $(\odot\mathrm{I}(L), \otimes, \to)$ is a residuated lattice.

Corollary 5.1. *Let L be a regular residuated lattice. Then $(\odot I(L), \subset)$ is a complete Heyting algebra (which is also called a frame).*

Proof. It is clear by Theorem 5.2 and Lemma 0-3.16 in [12]. In addition, we know indirectly that $(\odot\mathrm{I}(L), \subset)$ satisfies the infinite distributive law by Corollary 5.1.

Corollary 5.2. *For a regular residuated lattice L, $(\odot I(L), \subset)$ is an algebraic lattice.*

Proof. The notion of algebraic lattices, please refer to [12]. By Theorem 5.1 we know that $(\odot\mathrm{I}(L), \subset)$ is a complete lattice and the directed sups in $(\odot\mathrm{I}(L), \subset)$ are just set-unions. So it is easy to prove that the generated $\odot$-ideal $\langle a \rangle$ is a compact element of $(\odot\mathrm{I}(L), \subset)$ for any $a \in L$. Thus by Corollary 3.1 we have that for any $A \in \odot\mathrm{I}(L)$, $A = \bigcup\{\langle a \rangle | a \in A\}$ is a directed union of compact elements. This show that every element of $\odot\mathrm{I}(L)$ is a directed union of some compact elements. Thus $(\odot\mathrm{I}(L), \subset)$ is an algebraic lattice.

Acknowledgements. This work is supported by the NSF of China (10371106, 60774073).

References

1. Wang, G.: Non-classical Mathematical Logic and Approximate Reasoning. Science Press, Beijing (2003)
2. Wang, G., Zhou, H.: Introduction to Mathematical Logic and Resolution Principle, 2nd edn. Science Press, Beijing (2009)
3. Pavelka, J.: On fuzzy logic (I, II, III). Z. Math. Logic Grund. Math. 25, 45–52, 119–134, 447–464 (1979)
4. Wang, G.: MV-algebras, BL-algebras, R_0-algebras and multiple-valued logic. Fuzzy Systems and Mathematics 16(2), 1–15 (2002)
5. Xu, Y., Ruan, D., et al.: Lattice-valued logic. Springer, Berlin (2004)
6. Liu, L., Li, K.: Fuzzy filters of BL-algebras. Information Science 173, 141–154 (2005)
7. Liu, Y., Liu, S., Qin, K.: ILI-ideals and prime LI-ideals in lattice implication algebras. Information Science 155, 157–175 (2003)
8. Liu, C., Xu, L.: MP-filters of fuzzy implication algebras. Fuzzy Systems and Mathematics 23(2), 1–6 (2009)

9. Jun, Y., Roh, E., Xu, Y.: LI-ideals in lattice implication algebras. Bull. Korean Math. Sc. 35(1), 13–23 (1998)
10. Zhang, X., Xue, Z., Ma, Y.: Boolean MP-filters and Boolean MP-ideals in R_0-algebras (NM-algebras). Chinese J. of Engineering Mathematics 22(2), 287–294 (2005)
11. Wu, H., Shao, X.: A Method of Total Implication Reasoning on Complete Residuated Lattices. Advances in Mathematics (China) 35(3), 303–314 (2006)
12. Gierz, G., Hofmann, K.H., Keimel, K., et al.: Continuous Lattices and Domains. Cambridge University Press, Cambridge (2003)
13. Pei, D.: The characterization of resituated lattices and regular resituated lattices. Acta Mathematica Sinica 45(2), 271–278 (2005)

A Fuzzy System with Its Applications Based on MISS

Li-jiang Zhuang, Yu-bin Zhong*, Yan-qiang Li,
Cheng-han Wang, and Da-cheng Zeng

School of Mathematics and Information Sciences, Guangzhou University
Guangzhou 510006, P.R. China
e-mail: zhong_yb@163.com

Abstract. As the development of broadband network, E-government and informatization,the Mobile Intelligent Service System(*MISS*) is widely used in various fields. In practical applications, people find that the intelligence and security of MISS is very important. For the Asian Games will be held in Guangzhou in 2010, and the existing hotel service doesn't have unified service system, this paper designs and develops a fuzzy MISS .In this system, we use the fuzzy set to express data structures, use the fuzzy hierarchy synthetic evaluating model to assess hotel levels, and use the fuzzy cluster analysis method to provide each kind of tourists with different types of hotels . Finally ,we conduct it in evaluating and selecting hotels in tourism, and the result shows that the proposed system is secure, convenient and effective.

Keywords: Mobile intelligent service system,fuzzy set, fuzzy synthetic evaluation model, fuzzy cluster.

1 Introduction

In 2010, the Asian Games will be held in Guangzhou. This will be the greatest meeting that Guangzhou convenes throughout history, and this is also a very good opportunity for Guangzhou to promote its international reputation .Then, tourists coming from all over the world will swarm into Guangzhou, so the tourists' living life will be very important. But, the existing hotel service in Guangzhou doesn't have an unified system, tourists only be able to query the hotel and transact needed services separately, which is quite troublesome for external tourists who are not familiar with local hotel resources. According to Guangzhou existing service condition, tourists will dial 114 hot line to query the hotels for their communication methods or address, then make a

* Corresponding author.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 435–445.
springerlink.com

call to the corresponding hotel to query in details or transact services, and it will be quite inconvenient. What's worse, in the period of Asian Games, public hot lines as well as hotel hot lines may have a high possibility to become busier because tourists dial frequently. As a result, the consulting is even more difficult. So it is quite significant to study and develop a MISS for Guangzhou Asian Games(GAG),and build up a germ of urban resource overall planning system. This will enable tourists to connect to a public service platform to get information about all the hotels in Guangzhou in the shortest time. Now, we make a study of this problem.

2 Data Structures of Fuzzy MISS

To improve the running efficiency, this system takes logical tree as the data structures of hotel target system. The input and output stream are organized in linear table form during the communication between query sides and the server. The descriptions are shown in the flowing.

2.1 Logical Tree Data Structures

To accelerate the retrieval speed and improve inquire efficiency, we use tree data structure to organize and manage hotel target system (see Fig. 1).

2.2 Logical Linear Table Data Structures

In the course of communication between query clients and the server , the input and output data stream are organized in logical linear table ,which are split by using two separating characters: " : " and " * ". The information

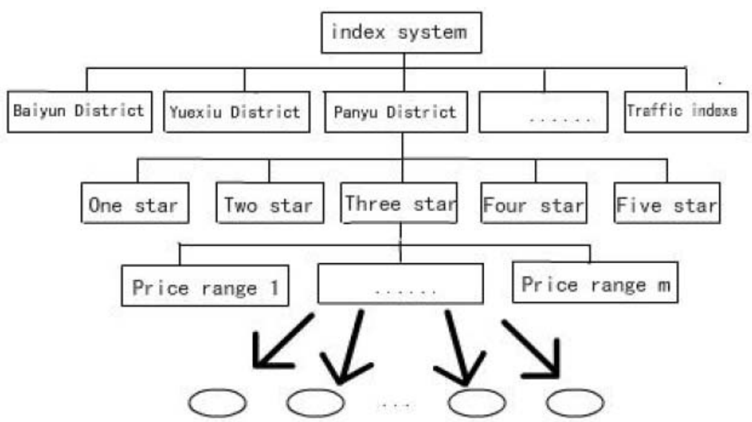


Fig. 1. Logical tree structure

Fig. 2. Logical linear table structure

is divided into elemental unit, in logic, these elemental units form linear table(Fig.2). Concretely, $MES * record_1 * record_2 * \cdots\cdots * record_n$ may be divided into Information[0], Information[1], ...Information [n-1].

According to fuzzy hierarchy synthetic evaluating model in reference [1], $B = A \circ R$. And here

$$A = \frac{a_1}{u_1} + \frac{a_2}{u_2} + \cdots + \frac{a_m}{u_m}(0 \leq a_i \leq 1)$$

is a fuzzy subset on factor field $U = \{u_1, u_2, \cdots\cdots, u_m\}$.

$$B = \frac{b_1}{v_1} + \frac{b_2}{v_2} + \cdots + \frac{b_m}{v_m}(0 \leq b_j \leq 1)$$

is a fuzzy subset on evaluating field $V = \{v_1, v_2, \cdots\cdots, v_m\}$. $R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ r_{m1} & r_{m2} & \cdot & r_{mn} \end{bmatrix}$ is a fuzzy relation between $U = \{u_1, u_2, \cdots\cdots, u_m\}$ and $V = \{v_1, v_2, \cdots\cdots, v_m\}$, and $B = A \circ R$ is the Fuzzy synthetic evaluating result. As for the problem of choosing hotel in Guangzhou, we analyzed each policy factor (here is target)on factor field and obtained the fuzzy subset on factor field.Then we established the fuzzy relation evaluating matrix R between the factor field and the evaluation field, and by $B = A \circ R$,we can obtain the requested hotel level of each kind of tourist, and use the fuzzy clustering method to provide different levels of hotels for each kind of tourist to choose from, so that it can help tourists make optimized choice while selecting hotels [2,3,4].

3 The Technical Architecture of the System

3.1 The Technical Architecture and Implemented Functions of the System

We chose Apache 2.2.10 for Web Server and *Ubuntu* 8.10 for Operate System. Apache is one of the most popular Web Server, it has fast and stable performance characteristics, and it can construct a server with quite stable and strong performance with Linux because of its perfect support for Linux. And we chose Java for core developing language. The Java version we adopted is the latest version of J2SE 6u11. We used *JSP* to develop user's Wap or Web query client and used Eclipse 3.4.1 to develop user's GUI, and chose

MySQL-5.0.22 for database server. Eclipse is today the most powerful Java integrated development environment (IDE), which integrated the function of prepare, compile, debug, running into one. And it has been in hot pursuit of numerous programmers, and Eclipse's *SWT* and *JFace* almost completely solved the problem that Java's *AWT* and Swing is in the inability to GUI development. So we use Eclipse to develop the hotel update client and the User Query Client. *MySQL* which has small size, high speed and low cost, has been widely used in the Java development, and *MySQL* is almost considered as the best partner in Eclipse's database application Development.

The System Architecture using hybrid structure with B/S model and C/S model, which is based on the decision to meet the demand characteristics of the various components of the system, it is flexible to satisfy the needs of the various parts. Because the performance of mobile phone can not compare with PC, the query procedures in the user query client must as simple as possible to meet the speed requirements. So we selected *JSP* for the development technology, and use web query to achieve the query function. And the hotel update client and Query Server and Management Server need a strong performance, so we used Java and Eclipse technology to develop. These technologies together build a powerful application platform.

Namely, we used the technology which makes up of Linux and Apache, *MySQL*, *JSP*, Java, Eclipse inthe system design process.

3.2 The Constructions of the System

This system is constructed by four parts (Fig. 3), and the details are shown as follows.

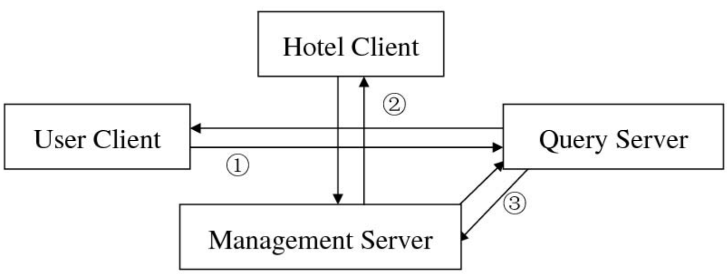


Fig. 3. The construction of the system

3.2.1 User Client

This part is developed with *JSP* that it is the technology which is selected to use based on the limited performance of mobile, with smart and fast response characteristics. When want to query hotel information and transact hotel business, Tourists can connect to the User Query Client by phone through

GPRS. Enter the relevant search keyword and click query button, Tourists will be able to get the required information, such as the hotel's location, near the traffic condition, distance away from the game site, hotel occupancy and favorable things, etc. And after logging in and passing validation, user can also directly transact hotel business like room booking and hotel taxi reservation through the query page.

3.2.2 Query Server

Query Server is developed with Java. When it is started, it is in a listening state all the time to wait for the user client requests. Once receiving request from user client, it will query the hotel information database which is connected by *JDBC-MySQL* driver with the user's keyword, and return the result set to user client in time. When user is using the query function, it is no need to logging in and verify, and the keyword and information will be sent to Query Server Expressly. But when user need to handle business, Query Server would demand user of logging in and verify user's password and personal information with password algorithm which is referenced above this section before allowing user to handle corresponding business. After user handle business, Query Server would communicate with Management Server to submit related information to Management Server, and wait hotel side to confirm and administer.

3.2.3 Hotel Client

This part is developed with Java and Eclipse technology. The Graphics User Interface (GUI) is developed with Eclipse. With this client, hotel user can manage Management Server's data and update the hotel information and occupancy information to let server keep the latest hotel information. In addition, hotel user can administer the reservation business user applied. Once the hotel side confirms the business, Hotel Client would send the administered information to Management Server, and tic the business item with complete mark from the queue waiting to be transacted. And hotel side can get the user's personal information according to information that user submitted. During this period, Hotel Client and Management Server will communicate, and return the result to the user. This process is not real-time, hotel side can not administer immediately after user submit the request. Hotel side is not likely to assign a worker to monitor the business through the Hotel Client in 24 hours. So, there will be a certain lag. It is believed that a good hotel would standardize the time to administer business in a certain time interval.

3.2.4 Management Server

This part is also developed with Java and Eclipse technology. It has friendly GUI. Management Server use multi-threading technology that allows multiple users to manage at the same time. But at the same time, it does not allow multiple user log on the same account and modify the same hotel's data for administer the same hotel's business. It is set to prevent confusion. Management Server will listen service port after started. Whenever receives a connection from hotel users, the server will open a new thread to verify the user's information, this process somewhat similar to handshake agreement. After verified, the hotel user can update data and process business through Management Server. Management Server can be set up to limit the number of users to log on to keep the server from network congestion and protect the server. Besides, the server has a log function to record event happening on the server running time. When a user tries to connect to the server, it will record automatically. When a user log on or log off, it will record the user's name, IP, log time. When the server is shut down, it will record the time. All of the log will save to the log file. Recording log is conducive to the management of maintenance and monitor the server's status, and unexpected events can be processed in a timely manner. After hotel user updates the hotel data, Management Server will communicate with Query Server in time to send the result updated to the Query Server, and let the tourist easy to get latest hotel information. In addition, after hotel user processed tourist's application, Management Server also communicates with Query Server to let it notifies the user about the application result of operation.

The 1st line in Fig.3 reflects the relation between User Client and Query Server. User Client send query request and business request to Query Server, and Query Server returns the query result and business result to User Client. They use Socket to communicate.

The 2nd line in Fig.3 shows the process of communication between Hotel Client and Management Server. Hotel Client send connection request information to Management Sever, the server opens a new connection thread after received the request and send the verification information to the client. Then the client sends name and password processed to the server to be verified in turn after received the above information. And then the server searches the database to match the user information received first. If exist, it would verify the password, otherwise interrupt their connection and destruct the thread for the client. If user's name is matched, it will enter the link of verifying password. The client processes password and sends it to the server to verify after received the request information for password. And the next is synchronous verification. The sever test whether multiple users are using the same account or whether multiple users are modifying the hotel information in the same time. If none of the above two conditions happening, the server will send authentication information to client, allow hotel user to update hotel information and administer business and every user's action will be recorded to the server log file.

The 3rd line in Fig.3 shows communication process between Management Server and Query Server. The two servers use and manage its own local databases. The database that Management Server uses is for account management and hotel management. And the database Query Server uses is hotel information database. Only when the tourist user submits application or hotel user administers business, these two databases transform their data. Remain time, they perform their respective roles. In the above three data transmission process, the information sender firstly transfers the information to cipher. And then package input stream and output stream, transfer information to UTF format string, so that the information has strong compatibility, receiver is not easy to receive garbled information.

4 The Security of Fuzzy MISS

The security of the proposed system is a major consideration [5]. To ensure the security of the hotel evaluation and selection process that leads to an instant booking of specific hotels for individual tourists, several specific features are designed in implementing a cryptographic algorithm in the proposed systems. These features include (a) determining the elliptic curve cryptographic procedure [6], (b) adopting space coordinates (X, Y, Z) for representing the points on elliptic curve, (c) defining CPK and IPA [7],(d) developing signature verification scheme [5], and finally (d) designing an encryption scheme [8]. These features are briefly discussed in the following.

The elliptic curve cryptographic system is based base on the elliptic curve discrete logarithm [9]. Some specific attack algorithms are adopted for the special type's elliptic curve discrete logarithm. For example, a SSSA attack algorithm is used for addressing an abnormal elliptic curve. The Pohlig-Hellman algorithm is used for calculating the elliptic curve discrete logarithm to attack if the rank of elliptic curves has no big prime number factor. To enhance the running speed of mobile platform, the widely used famous NID_X9_62_prime192v1 elliptic curve parameter group in Open SSL is adopted.

This system use CPK2.0 compose public key infrastructure. It has several advantages including (a) ability to add a system random factor as random private key for avoiding the user collusion cracking potential safety hazard of private key matrix. (b) use of a CA storage center The use of the system is different from the existing mode of public-key systems Public Key Infrastructure. A new public-key management model based on Identity-based Public-key Authentication technology (IPA) is adopted for overcoming the shortcomings of PKI which used in mobile electronic terminal.

Digital signature is a signature method to a message in electronic form. It's some number of strings that any other person can not be forged. This special series digital is an authenticity to prove the signature. In the electronic information transmission process, the adoption of digital signatures can achieve the same effect as traditional handwritten signature. To achieve compatibility with the PC platform, especially for compatible with the famous OpenSSL, the Elliptic Curve Digital Signature Algorithm (ECDSA) [8] is used.

An encryption scheme using digital envelope technology is developed in the proposed system for making the system run faster and secure. Digital envelope is a practice application of public-key cryptosystem. It uses encryption technology to ensure that only the specific recipient can read the contents of communications with high security [10-12].

5 The Implementing Scheme of Fuzzy MISS

In this system, we use the comparative concrete Fuzzy Set to make knowledge representation of the hotel information and use the Fuzzy Hierarchy Synthetic Evaluating Model to assess hotel level of each kind of traveler, finally we use the Fuzzy Cluster Analysis method to provide each kind of traveler with different types of hotels to choose. This system is simple, fast, reliable and secure; users can easily query and transact hotel businesses by using telephones or other communication facilities. So, it plays a positive role in solving the tourist's traveling life problems, and it promotes the development of Guangzhou electronics and information industry as well as the M-Commerce. So,we developed the Mobile Intelligent Hotel Service System.

5.1 Client Connects Server

When client apples for query request, system writes $CLIENT : +IPADDRESS+KEYWORDS$into the input stream automatically . And $IPADDRESS$ is client's IP address; $KEYWORDS$ is composed of two parts, and the form is:

$$operationprompt : SQLlanguage : .$$

And here, *operation prompt* expresses the type of the request which is put forward by client. All of the operation prompts established in the system are as follows: (see Table 1)

Table 1. *Operations*

SEND	Users send message to hotel
SELECTHOTEL	Search hotels and complete the simple search at first
SELECTCLIENT	Search orders and return the order information of clients
ORDER	Order hotel and create the order
QUIT	break off the connection

The SQL language is created automatically according to user's operation by the system. Each part of the information which is written into the input stream is separated by ":", so that it is convenient for server procedure to separate and process.

5.2 Server Returns to Client

After accepting the request of client, server make corresponding processing according to the information which is written in the input stream by client , and return the record set after retrieving the database . The server returns records to client in a specific form , and the concrete form are as follows:

$$MES * record_1 * record_2 * \cdots\cdots * record_n,$$

and here $MES*$ is information head, $record_i$ is information of record i which is inquired according to the user's request. The form of $record_i$ is: $Data_1 : Data_2 : Data_3 : \ldots\ldots Data_m$, and here $Data_i$ returns the concrete information of a certain aspect of hotels or orders.

Each part is separated by " : ", therefore it is convenient for client procedure to separate with $nextToken()$, then system uses friendly, natural, and beautiful interface to present the inquired information.

5.3 The Data Stream between Inquiry Client and Server

The data stream in Query Server and between servers is using the multi-threaded concurrent collaboration and Socket programming in order to achieve efficient transmission.

(1) Multi-threaded Concurrent Collaboration

Multi-thread programming can cause the procedure having two or more than two concurrent implementation of clues, such as multiple people co-operate to complete a task in daily work, which can improve procedures in response performance and increase the efficiency of resource use in many cases. In this system, the Management Server is on the use of multi-threading technology. When a user applies to connect the server, the server will create a thread to execute verification steps in order and allow user to operate the data after user passed the verification. This practice will enable the server has the ability to process the number of request and improve the server's operational efficiency.

(2) Socket Programming

The core of the traditional C/S model network program is that transferring information between client and server through the network connection. Sometimes the transferring data is called message. And *TCP* Socket connection is generally used between client and server. The following is the communication of client and server (see Fig 4).

The specific process of communication are as follows:

1) Server start monitoring procedures, listen to the specified monitor port, and wait to receive client connection requests.

2) Client program starts and requests to connect with the server specified port.

3) Both client and server open two streams after a successful connection, then the client input stream is connected to the server output stream, and

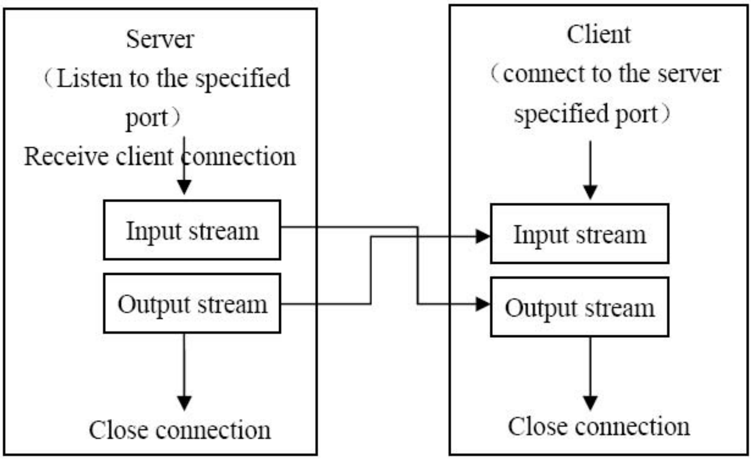


Fig. 4. Communication model

the server input stream is connected to the client output stream. They can have two-way communication after the connection of two sides' stream has established.

4) The connection between client and server will be disconnected after communication was completed.

6 Conclusion

1. The system can solve the problem of tourists' living life during Asian Games to a great extent, which enable tourists to query relate information easily and make ideal decisions. Therefore, it is a true *MISS* system.

2. This system uses cipher algorithm, so, the software are encrypted. Consequently, it is a true safe system.

3. In view of the certain limitation of realizing encryption and decryption on *PC* just using software, specially the slow processing speed on *PC*, it has not been able to meet the application need well in the large-scale system. so this plan selects the hardware method, in application, it may obtain fast performance speed, thus enhances overall system's performance.

4. Further study on this system can be done, such as improving the efficiency of Elliptic Curve algorithm, developing more useful IC card-*CPU* card, to improve this *MISS* system to make it a set of very popular *MISS* which holds the suitable share in the market.

5. This system belongs to a new development direction of electronic commerce— M-Commerce. It has enormous market prospect and development potential in the development of the electronic commerce economy in our country . This project makes great contributes to promoting the development of our country's informatization. If it can get further study, it

is sure that our country's informatization construction will be taken to a new stair. At the same time,it will also solve the problems of GAG in 2010.

Acknowledgements. This research is supported by the Foundation of the Major Innovation Experiments Projects of Undergraduates of the Education Department of Guangdong Province(No.1107810051).

References

1. Zhong, Y.: The FHSE Model of Software System for Synthetic Evaluating Enterprising. Journal of Guangzhou University 4(4), 316–320 (2005)
2. Zhong, Y.: The Design and Research of Controller in Fuzzy PETRI NET. Advance in Soft Computer (40), 41–49 (2007)
3. Zhong, Y.: The Optimization for Location for Large Commodity's Regional Distribution Center. Advance in Soft Computer (40), 969–979 (2007)
4. Houwen, Z., Canlong, L.: One kind of elliptic curve password card's design with realizes. Micro Computer Information l22, 1–3 (2006)
5. Tang, C.: Non-interactive ID-based UDVS Proof From Bilinear Pairings. Journal of Guangzhou University 7(3), 13–21 (2008)
6. Zhong, Y.: The Design and Research of Controller in Fuzzy PETRI NET. Advance in Soft Computer 40, 41–49 (2007)
7. Ribeiro, R.A.: Fuzzy multiple attribute decision making: A review and new preference elicitation techniques. Fuzzy Sets and Systems 78, 155–181 (1996)
8. Dong, J.: Implement of Elliptic Curve Cryptography on Singlechip. Financial Computer of Huanan 3, 36–38 (2003)
9. Zhong, Y.: The Optimization for Location for Large Commodity's RD Center. Advance in Soft Computer 40, 969–979 (2007)
10. Yeh, C.H., Deng, H.: A Practical Approach to Fuzzy Utilities Comparison in Fuzzy Multicriteria Analysis. International Journal of Approximate Reasoning 35(2), 179–194 (2004)
11. Zhong, Y., Mi, H., Chen, H.: Design and Optimization Algorithm of VCS under Security Protection. In: IEEE of International Conference on Web Information Systems and Mining, pp. 552–556 (2009)
12. Zhong, Y., Deng, H., Chen, H.: A Fuzzy Logic Based Mobile Intelligent System for Effectively Evaluating and Selecting hotels in Tourism. In: IEEE of International Conference on Web Information Systems and Mining, pp. 733–737 (2009)

Particle Swarm Optimization of T-S Fuzzy Model

Ming Ma[1], Li-biao Zhang[2], and Yan Sun[1,2]

[1] College of Mathematics, Beihua University, Jilin 132013, china
e-mail: ma9063@163.com
[2] College of Computer Science and Technology, Jilin University, Changchun 130012, China
e-mail: mam@mail.edu.cn

Abstract. This paper introduces a new algorithm for Takagi -Sugeno (T-S) fuzzy modeling based on particle swarm optimization. Compared the standard particle swarm optimization, in the proposed algorithm a mixed code is adopted to represent a solution. Binary codes represent the structure of T-S fuzzy model, and real values represent corresponding parameters. Numerical simulations show the effectiveness of the proposed algorithm.

Keywords: Particle swarm optimization; T-S fuzzy model; fuzzy rules.

1 Introduction

Fuzzy systems have become an active research area in recent years. Several fuzzy model designs have been proposed including knowledge-driven modeling and data-driven modeling. Compared to knowledge-driven modeling, data-driven fuzzy modeling has been applied to many fields, such as pattern recognition, data mining, classification, prediction, and process control, and etc [1-5]. Generally speaking, the so-called data-driven fuzzy modeling is just an optimization process in determining the structure and the parameter of a fuzzy system via sample data.

Fuzzy system identification is one of the main approaches of fuzzy system modeling. The accuracy of fuzzy system model relate to the result of fuzzy system identification. In this paper, an algorithm based on Particle Swarm Optimization (PSO) is proposed to optimize Takagi-Sugeno (T-S) model, in the proposed algorithm we adopted a binary value vector and a real values vector to represent a solution, and used the different equation to update the different parameters. Satisfactory results through experiments are obtained.

The rest of this paper is organized as follows: The T-S model is introduced in Section 2. The proposed algorithm is described in Section 3. The

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 447–452.
springerlink.com

simulation and experimental results are presented in Section 4. Finally, concluding remarks are given in Section 5.

2 T-S Model

Takagi-Sugeno (T-S) model is a fuzzy system proposed by Takagi and Sugeno in 1985[6]. As a method of data-driven modeling, it has been successfully used in a wide variety of applications. In the model the ith fuzzy rule have the form

$$\mathbf{R_i:} \; x_1 \; is \; A_{i1}, ..., x_n \; is \; A_{in} \; then \; y_i = \; c_{i0} + c_{i1}x_{i1} + ... + c_{in}x_n \tag{1}$$

where n is the number of input variables.$i = 1 \ldots r$,and r is the number of if-then rules. A_{ij} is the antecedent fuzzy set of the ith rule. y_i is the consequence of the ith if-then rule .$c_{ij}(i = 1 \ldots r; j = 1 \ldots n)$ is real number. Then by using center of gravity method for defuzzification, we can represent the T-S system as:

$$\mathbf{y} = \frac{\sum_{i=1}^{r} y_i \prod_{j=1}^{n} \mu_{Aij}(x_i)}{\sum_{i=1}^{r} \prod_{j=1}^{n} \mu_{Aij}(x_i)}. \tag{2}$$

3 Pruning Algorithm

3.1 PSO

Particle Swarm Optimization (PSO) is an optimization algorithm proposed by Kennedy and Eberhart in 1995 [7,8]. It is easy to be understood and realized and has been applied in many optimization problems [9-11]. PSO originated from the research of food hunting behaviors of birds. Each swarm of PSO can be considered as a point in the solution space. If the scale of swarm is N, then the position of the $i-th(i = 1, 2 \ldots N)$ particle is expressed as X_i . The "best" position passed by the particle is expressed as pBest[i]. The speed is expressed with V_i. The index of the position of the "best" particle of the swarm is expressed with g. Therefore, swarm i will update it's own speed and position according to the following equations:

$$V_i = wV_i + c_1 rand()(pbest[i] - X_i) + c_2 Rand()(pbest[g] - X_i), \tag{3}$$

$$X_i = X_i + V_i, \tag{4}$$

where c_1 and c_2 are two positive constants, $rand()$ and $Rand()$ are two random numbers within the range [0, 1], and w is the inertia weight. The equations consist of three parts. The first part is the former speed of the swarm, which shows the present state of the swarm; the second part is the cognition modal, which expresses the thought is the cognition modal, which expresses the thought of the swarm itself; the third part is the social modal. The three parts together determine the space searching ability. The first part has the ability to balance the whole and search a local part. The second part causes the swarm to have a strong ability to search the whole and avoid local minimum. The third part reflects the information sharing among the swarms. Under the influence of the three parts, the swarm can reach an effective and best position.

3.2 Description of the Algorithm

The pruning algorithm based on PSO is formed of two phases. In the first phase, we use n dimensional real-valued vector to represent a solution, as shown in follows:

$$X = (x_1, x_2, ..., x_n), \tag{5}$$

where n is the number of all the real-valued parameters, it include the parameters of membership functions and consequence parameters $c_{ij}(i = 1 \ldots r; j = 1 \ldots n)$.

The following functions have been used for evaluation of PSO.

$$F(x) = \frac{1}{\sum\limits_{K}(O - T)^2}, \tag{6}$$

where K is the number of sample, T is the teacher signal, and O is the output.

The real-valued parameters update it's own speed and position according to the PSO. Compared with the consequence parameter, the optimization of membership functions only adjusts a little; therefore, the parameters of membership functions update it's speed and position according to the following equations:

$$V_i = wV_i + c_1 rand()(pbest[i] - X_i) + c_2 Rand()(pbest[g] - X_i), \tag{7}$$

$$X_i = X_i + c_3 G(V_i), \tag{8}$$

where c_3 is a positive constant, $c_3 << 1$,function $G(x)$ showed as follows:

$$F(x) = \frac{1 - exp(-x)}{1 + exp(-x)}. \tag{9}$$

Thus if $x > 0$, then $c_3 * G(X)$is in (0,1) otherwise $c_3 * G(X)$ is in (-1,0).

If the fixed precision is achieved, then go to the second phase. In the second phase, we use a mixed vector to represent a solution, as shown in follows:

$$X = (e_1, e_2, ..., e_r, x_1, x_2, ..., x_m), \tag{10}$$

where r is numbers of fuzzy rules, and $e_i(i = 1 \ldots r)$ is 0 or 1, if the value of the e_i is 1, then the corresponding fuzzy rule is enabled, otherwise it is disabled;

m is numbers of consequence parameters, and $x_i(i = 1 \ldots m)$ represents the corresponding $c_{ij}(i = 1 \ldots r; j = 1 \ldots n)$.

First random select a e_i and update it according to the discrete PSO [12] . In addition, when the value of the e_i is 0, that is to say, the rule is redundant, the corresponding x_i is disabled, so consequence parameters $x_i(i = 1 \ldots m)$ update it's speed and position according to the following equations:

$$V_{id} = e_{id}(wV_{id} + c_1 rand()(pbest[id] - X_{id}) + c_2 Rand()(pbest[gd] - X_{id})), \quad (11)$$

$$X_{id} = X_{id} + V_{id}. \quad (12)$$

If the fixed precision is achieved, then stop and cut the fuzzy rule that the corresponding e_i is 0. Initialize the particle swarm again, and loop the process until it achieves the termination condition, we will obtain a near-optimal structure of T-S model in the end.

3.3 The Execution of the Algorithm

The algorithm:

1. Initialize the particle swarm: Designate the population size N, generate speed Vi and position Xi of each particle randomly, and let k=0; Evaluate the fitness of each particle.
2. Update the parameters of membership functions and the other real-valued parameters.
3. Evaluate the fitness of each particle.
4. If the fixed precision is achieved then go to 5 otherwise go to 2.
5. Random select a e_i to update.
6. Update the binary parameters e_i of each particle according to the discrete PSO .
7. Update consequence parameters.
8. If the pruning condition is achieved then go to 9, otherwise go to 6.
9. Pick the best particle to cut the redundant fuzzy rule.
10. If the termination condition is achieved then stop, otherwise go to 5.

4 Numerical Simulations

We used the proposed algorithm to the Mackey-Glass time series, which is generated by the following time-delay differential equation [13]:

$$\frac{dx(t)}{dt} = \frac{0.2x(t-17)}{1 + x^{10}(t-17)}. \quad (13)$$

In our simulations we have considered $x(t-1), x(t-2)$and$x(t-3)$ as inputs to predict $x(t)$.we set the initial condition is: $x(0) = 0.64$ and generated a sample of 1000 points, The first 500 points were used as training data, and the last 500 points as test data to validate the model's performance.

In our algorithm we used 3 fuzzy sets for each input, and the Gaussian membership function is used for each fuzzy subset. The parameters of the PSO are these: learning rate $c_1 = c_2 = 2, c_3 = 0.005$, inertia weight is taken from 0.9 to 0.2 with a linear decreasing rate. The population of particles was set to 60.Before the execution of the algorithm the T-S model has 27 fuzzy rules. The comparative results between the model in this paper and other models are summarized in Table 1.

Table 1. Comparative results

Model	Number of inputs	Rules	RMSE
T-Sekouras[13]	4	6	0.0041
ANFIS [14]	4	16	0.0016
KUKOLJ[15]	4	9	0.0061
WANG[16]	9	121	0.01
OUR MODEL	3	12	0.0034

The results in table1 have proved that the proposed pruning algorithm is applicable and efficient.

5 Conclusion

Fuzzy system identification includes structure identification and parameters identification. Based on the standard PSO, the proposed pruning algorithm can obtain a near-optimal T-S model. The numerical experiments indicate the effectiveness of the algorithm.

References

1. Jang, J.R.: ANFIS: adaptive network-based fuzzy inference system. IEEE Trans. on Systems, Man Cybernetics 23, 665–685 (1999)
2. Kim, E., Park, M., Ji, S., Park, M.: A new approach to fuzzy modeling. IEEE Transactions on Fuzzy Systems 5(3), 328–337 (1997)
3. Setnes, M., Roubos, H.: GA-fuzzy modeling and classification: complexity and performance. IEEE Trans. Fuzzy Systems 8(5), 509–522 (2000)
4. Xing, Z.Y., Jia, L.M., Yong, Z.: A Case study of data-driven interpretable fuzzy modeling. ACTA Automatica SiT1ICA 31(6), 815–824 (2005)
5. Mastorocostas, P.A., Theocharis, J.B., Petridis, V.S.: A constrained orthogonal least-squares method for generating T-SK fuzzy models: application to short-term load forecasting. Fuzzy Sets and Systems 118(2), 215–233 (2001)
6. Takagi, T., Sugeno, M.: Fuzzy identification of system s and its application to modeling and control. IEEE Trans. on Systems, Man and Cybernetics 15(1), 116–132 (1985)

7. Eberhart, R., Kennedy, J.: A new optimizer using particle swarm theory. In: Proc. 6th Int. Symposium on Micro Machine and Human Science, pp. 39–43 (1995)
8. Kennedy, J., Eberhart, R.: Particle Swarm Optimization. In: IEEE International Conference on Neural Networks IV, pp. 1942–1948 (1995)
9. Eberhart, R., Hu, X.: Human tremor analysis using particle swarm optimization. In: Proceedings of the IEEE Congress on evolutionary computation, pp. 1927–1930 (1999)
10. Parsopoulos, K.E., Vrahatis, M.N.: Particle Swarm Optimizer in Noisy and Continuously Changing Environments. Artificial Intelligence and Soft Computing, 289–294 (2001)
11. Parsopoulos, K.E., Vrahatis, M.N.: Particle swarm optimization method for constrained optimization problems. In: Proceedings of the Euro-International Symposium on Computational Intelligence
12. Kennedy, J., Eberhart, R.C.: A discrete binary version of the particle swarm algorithm. In: Proc. on Systems, Man, and Cybernetics (1997)
13. T-Sekouras, G., Sarimveis, H., Kavakli, E., Bafas, G.: A hierarchical fuzzy-clustering approach to fuzzy modeling. Fuzzy Sets and Systems 150, 245–266

Inverse Limits of Category CL

Hai-yang Li

School of Science, Xi'an Polytechnic University, 710048, Xi'an, P.R. China
e-mail: fplihaiyang@126.com

Abstract. In this paper, the closed-set-latticefication of join-semilattices, a new approach that the join-semilattice becomes the closed-set lattice, is introduced. Moreover, by means of it, the structure of inverse limits in **CL**, the category of closed-set lattices, is given.

Keywords: Join-semilattices; Closed-set lattices; Closed-set-latticefication; Category; Inverse limits.

1 Definitions and Preliminaries

Let L be a complete lattice, and $Copr(L)$ the set of nonzero co-prime elements of L. If $Copr(L)$ is a $\vee$-generating set of L (i.e. for each $x \in L$, there exists a subset $B_x \subset Copr(L)$ such that $x = \bigvee B_x$), then we call L a closed-set lattice. It can be proved ([2,8]) that a complete lattice L is a closed-set lattice if and only if it is isomorphic to the lattice $(\mathcal{J}', \subset)$ of closed sets of a topological space $(X, \mathcal{J})$ ($\mathcal{J}' = \{X - V \mid V \in \mathcal{J}\}$ is called a closed topology on X), or equivalently, L^{op} is a spatial locale [7]. The category of closed-set lattices and mappings which preserves arbitrary unions and nonzero co-prime elements is denoted by **CL**.

Limits theory is an important part in category theory, and inverse limits and direct limits play an important role in limits theory. Therefore the structures of inverse limits and direct limits have caused interests widely. For example, inverse limits of the categories of locales **Loc** and topological molecular lattice **TML** were studied ([3,11]), direct limits of the categories of spatial locales **SLoc** and meet-continuous lattice were discussed ([4, 6]), and the limits structure of molecular lattice category was given ([10]). In [9], we have studied some categorical properties of **CL**. As the continuation of [9], we mainly study the structures of inverse limits of category **CL** in this paper.

We define a relation $\lhd$ on closed-set lattices $(L, \leq)$ by putting

$$a \lhd b \Longleftrightarrow a \leq c \leq b \ \text{ for some } c \in Copr(L).$$

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 453–458.
springerlink.com

Then it can be easily verified that $\lhd$ is approximating and satisfies interpolation property, and $x \lhd x$ if and only if $x \in Copr(L)$.

Lemma 1. [9] *Suppose that L_1 and L_2 are closed-set lattices, and $f : L_1 \longrightarrow L_2$ a $\vee$-preserving mapping (consequently, f has a right adjoint $f^* : L_2 \longrightarrow L_1$, see [2]). Then the following statements are equivalent:*
(1) f is a **CL**-*morphism;*
(2) f preserves relation $\lhd$;
(3) f^* preserves finite unions (particularly, $f^*(0) = 0$).

2 The Closed-Set-Latticefies of Join-Semilattice

Let L be a join-semilattice. We write $\Downarrow a = \{b \in L^* \mid b \leq a\}$, where $L^* = Copr(L)$. Let $\mathcal{B} = \{\cup_{a\in F} \Downarrow a \mid F$ is a finite subset of $L\}$ and $\widetilde{L}$ be the family of all lower sets A of $(L^*, \leq)$ satisfying:
($*$) If $a \in L^*$ and $A \not\subset P$ (for each $P \in \mathcal{B}(a)$), then $a \in A$,
where $\mathcal{B}(a) = \{P \in \mathcal{B} \mid a \notin P\}$.

It is easy to verify that $\Downarrow a \in \widetilde{L}$ for every $a \in L$.

Lemma 2. $(\widetilde{L}, \subset)$ *is a complete lattice, and for any family $\{A_j\}_{j\in J} \subset \widetilde{L}$, $\bigwedge_{j\in J} A_j = \bigcap_{j\in J} A_j$ and $\bigvee_{j\in J} A_j \overset{(**)}{=} \{a \in L^* \mid \forall P \in \mathcal{B}(a), \bigcup_{j\in J} A_j \not\subset P\}$.*

Proof. Let $B = \bigcap_{j\in J} A_j$ and $a \in L^*$. If, for each $P \in \mathcal{B}(a), B \not\subset P$, then $A_j \not\subset P$ for every $j \in J$. Since A_j ($\forall j \in J$) satisfies condition ($*$), $a \in A_j$ which implies that $a \in B$. Therefore B satisfies condition ($*$), and it follows that $B \in \widetilde{L}$ and $\bigwedge_{j\in J} A_j = B$. As L^* is the greatest element of $(\widetilde{L}, \subset)$, $(\widetilde{L}, \subset)$ is a complete lattice.

Denote the right of equality ($**$) by A. Obviously, A is a lower set in $(L^*, \leq)$. Next, suppose that $a \in L^*$ and $A \not\subset P$ for every $P \in \mathcal{B}(a)$. We will show that $a \in A$. Assume that $a \notin A$. By the definition of A, $\bigcup_{j\in J} A_j \subset P_0$ for some $P_0 \in \mathcal{B}(a)$. As $A \not\subset P_0$, there exists a $y \in A - P_0$. We have $\bigcup_{j\in J} A_j \not\subset P_0$ by the definition of A again. This is a contradiction. Hence $a \in A$, which means that A satisfies condition ($*$) and thus $A \in \widetilde{L}$. Finally, suppose that $B \in \widetilde{L}$ satisfying $A_j \subset B$ for each $j \in J$. We will show that $A \subset B$ (accordingly, $A = \bigvee_{j\in J} A_j$). Since $B \in \widetilde{L}$, it suffices by condition ($*$) to show that $B \not\subset P$ for each $a \in A$ and each $P \in \mathcal{B}(a)$. By the definition of A, $\bigcup_{j\in J} A_j \not\subset P$, particularly $B \not\subset P$. This completes the proof of Lemma 2.

Lemma 3. $(\widetilde{L}, \subset)$ *is a distributive lattice.*

Proof. Let $A,\ B,\ C \in \widetilde{L}$ and $x \in A \cap (B \vee C)$. It suffices to show $x \in B \cup C$. Assume that $x \notin B$ and $x \notin C$. As B and C satisfy condition ($*$), there exist $P_1 \in \mathcal{B}(x)$ and $P_2 \in \mathcal{B}(x)$ such that $B \subset P_1$ and $C \subset P_2$. Let $P = P_1 \cup P_2$. Then $P \in \mathcal{B}(x)$ and $B \cup C \subset P$. By the definition of $B \vee C$, $x \notin B \vee C$. This is a contradiction.

Corollary 1. *If $\mathcal{A}$ is a finite subset of $\widetilde{L}$, then $\bigvee \mathcal{A} = \bigcup \mathcal{A}$.*

Lemma 4. *Denote $N = \{\Downarrow a \mid a \in L^*\}$, then $N \subset Copr(\widetilde{L})$ and N is the $\vee$-generating set of $\widetilde{L}$.*

Proof. Suppose that $\Downarrow a \subset A_1 \vee A_2$ but $\Downarrow a \not\subset A_1$ and $\Downarrow a \not\subset A_2$, where $a \in L^*$ and $A_1,\ A_2 \in \widetilde{L}$. Therefore $a \notin A_1$ and $a \notin A_2$ which implies that $a \notin A_1 \cup A_2 = A_1 \vee A_2$. This contradicts the $\Downarrow a \subset A_1 \vee A_2$. Then $\Downarrow a \in Copr(\widetilde{L})$. For every $P \in \widetilde{L}$, we have $\bigvee_{a\in P} \Downarrow a \subset P = \bigcup_{a\in P} \Downarrow a \subset \bigvee_{a \in P} \Downarrow a$. Hence N is the $\vee$-generating set of $\widetilde{L}$.

By the definition of closed-set lattice and above lemmas, we have the following theorem:

Theorem 1. $(\widetilde{L}, \subset)$ *is a closed-set lattice.*

We call that $(\widetilde{L}, \subset)$ is the closed-set-latticefies of join-semilattice $(L, \leq)$.

3 Inverse Limits of Category CL

We refer to [1] for some categorical notions.

Let $\{A_i,\ f_{ij},\ I\}$ be an inverse system of closed-set lattices. We denote by $\prod_{i\in I} A_i$ the direct products of $\{A_i \mid i \in I\}$. Let $A = \{x = \{x_i\}_{i\in I} \in \prod_{i\in I} A_i \mid \forall i, j \in I, i \leq j$, we have $f_{ij}(x_i) = x_j\}$. Then, for every $x \in A$, $\pi_j(x) = f_{ij} \circ \pi_i(x)$, where $i \leq j$ and π_i is a projection.

It is easy to verify the following lemma:

Lemma 5. $(A, \leq)$ *is a complete lattice, and for any family $\{x^s\}_{s\in S} \subset A$, $\bigvee_{A_{s\in S}} x^s = \bigvee_{s\in S} x^s, \bigwedge_{A_{s\in S}} x^s = \bigvee\{t \mid t \leq x^s$ and $t \in A\}$, where $\leq$ is the point-wise order, $\bigvee_A$ is the unions in A and $\bigvee$ is the unions in $\prod_{i\in I} A_i$.*

Obviously, $h_i = \pi_i|A : A \longrightarrow A_i$ preserves arbitrary unions, and thus $h_i^* : A_i \longrightarrow A$ preserves arbitrary intersections, where h_i^* is the right adjoint of h_i.

By the closed-set-latticefies of join-semilattice A, we have the following theorem:

Theorem 2. $(\widetilde{A}, \subset)$ *is a closed-set lattice, and for any family $\{P_j\}_{j\in J} \subset \widetilde{A}$, $\bigwedge_{j\in J} P_j = \bigcap_{j\in J} P_j$ and $\bigvee_{j\in J} P_j = \{a = \{a_i\}_{i\in I} \in A^* \mid \forall P \in \mathcal{B}(a), \bigcup_{j\in J} P_j \not\subset P\}$, where $A^* = \{a = \{a_i\}_{i\in I} \mid a \in A$ and $a_i \in Copr(A_i)\}$, $\Downarrow a = \{b \mid b \leq a$ and $b \in A^*\}$ $(a \in A)$, and $\mathcal{B} = \{\cup_{a\in F} \Downarrow a \mid F$ is the finite subset of $A\}$. $\widetilde{A}$ is the family of all lower sets of $(A^*, \leq)$ satisfying*

$(*)$ If $a \in A^*$ and $B \not\subset P$ (for each $P \in \mathcal{B}(a)$), then $a \in B$, where $\leq$ is the point-wise order and $\mathcal{B}(a) = \{P \in \mathcal{B} \mid a \notin P\}$.

If $A^* = \emptyset$, then $\widetilde{A} = \{\emptyset\}$. It is easy to verify the following theorem:

Theorem 3. *$\{\widetilde{A}, p_i|_{i\in I}\}$ is the inverse limits of the inverse system $\{A_i,\ f_{ij},\ I\}$, where $p_i : \widetilde{A} \longrightarrow A_i$ is defined by $p_i(\emptyset) = 0_{A_i}$ $(\forall i \in I)$.*

In the following, we will consider the condition of $A^* \neq \emptyset$.

Lemma 6. *Let $g : A \longrightarrow \widetilde{A}$ is defined by $g(a) = \Downarrow a$ $(\forall a \in A)$, then g preserves arbitrary intersections.*

Proof. For any family $\{a^s\}_{s\in S} \subset A$, it is easy to see $\Downarrow (\bigwedge_{A_{s\in S}} a^s) \subset \bigcap_{s\in S} \Downarrow a^s$. Conversely, if $a \in \bigcap_{s\in S} \Downarrow a^s$, then $a \leq a^s$ for each $s \in S$. By the definition of $\bigwedge_A$, we have $a \leq \bigwedge_{A_{s\in S}} a^s$. So $a \in \Downarrow (\bigwedge_{A_{s\in S}} a^s)$, and thus $\bigcap_{s\in S} \Downarrow a^s = \Downarrow (\bigwedge_{A_{s\in S}} a^s)$. This completes the proof of Lemma 6.

Theorem 4. *$\{\widetilde{A},\ p_i|_{i\in I}\}$ is the inverse limits of the inverse system $\{A_i,\ f_{ij},\ I\}$, where $p_i = h_i \circ g_*$ and g_* is the left adjoint of g.*

Proof. **Step 1.** $\{\widetilde{A},\ p_i|_{i\in I}\}$ is the natural source of the inverse system $\{A_i,\ f_{ij},\ I\}$.

Firstly, p_i is a **CL**-morphism. Since $h_i,\ g_*$ preserve arbitrary unions, p_i $(i \in I)$ preserves arbitrary unions. For any family $\{b_t\}_{t\in T} \subset A_i$, $\Downarrow h_i^*(\vee_{t\in T} b_t) = \vee_{t\in T} \Downarrow h_i^*(b_t)$, where T is the finite subset. In fact, if $x \in \Downarrow h_i^*(\vee_{t\in T} b_t)$, then $x \in A^*$ and $x \leq h_i^*(\vee_{t\in T} b_t)$. Hence $h_i(x) \leq h_i \circ h_i^*(\vee_{t\in T} b_t) \leq \vee_{t\in T} b_t$, and thus exists $t \in T$ such that $h_i(x) \leq b_t$. This implies $x \leq h_i^*(b_t)$ and $x \in \vee_{t\in T} \Downarrow h_i^*(b_t)$. Obviously, $\Downarrow h_i^*(\vee_{t\in T} b_t) \supset \vee_{t\in T} \Downarrow h_i^*(b_t)$, which means $\Downarrow h_i^*(\vee_{t\in T} b_t) = \vee_{t\in T} \Downarrow h_i^*(b_t)$. This follows that $p_i^*(\vee_{t\in T} b_t) = g \circ h_i^*(\vee_{t\in T} b_t) = \Downarrow h_i^*(\vee_{t\in T} b_t) = \vee_{t\in T} \Downarrow h_i^*(b_t) = \vee_{t\in T} p_i^*(b_t)$. Hence p_i^* preserves finite unions.

Secondly, $p_j = f_{ij} \circ p_j$ $(i \leq j)$. In fact, $f_{ij} \circ p_i = f_{ij} \circ h_i \circ g_* = h_j \circ g_* = p_j$.

Step 2. Assume that $\{B,\ q_i|_{i\in I}\}$ is also a natural source of the inverse system $\{A_i,\ f_{ij},\ I\}$, we will prove that there exists a unique **CL**-morphism $f : B \longrightarrow \widetilde{A}$ such that $q_i = p_i \circ f$ $(\forall i \in I)$.

Existence. As $\{B,\ q_i|_{i\in I}\}$ is the natural source of the inverse system $\{A_i,\ f_{ij},\ I\}$, $f_{ij} \circ q_i(b) = q_j(b)$ $(b \in B,\ i, j \in I$, and $i \leq j)$. It implies that $\{q_i(b)\}_{i\in I} \in A$. Let $f'(b) = \Downarrow \{q_i(b)\}_{i\in I}$, then $f' : B \longrightarrow \widetilde{A}$ is a mapping. Again, let $f(a) = \bigvee_{s\in S} f'(s) = \bigvee_{s\in S} \Downarrow \{q_i(b)\}_{i\in I}$, where $S = \{s \mid s \leq a$ and $s \in Copr(B)\}$. Then $f : B \longrightarrow \widetilde{A}$ is a mapping.

Firstly, $q_i = p_i \circ f$. For each $s \in S$, $g_*(\Downarrow \{q_i(s)\}_{i\in I}) = \{q_i(s)\}_{i\in I}$ by $g_*(\Downarrow \{q_i(s)\}_{i\in I}) = \bigwedge_A \{c \in A \mid \Downarrow \{q_i(s)\}_{i\in I} \subset \Downarrow c\}$ and $\{q_i(s)\}_{i\in I} \in A^*$. Hence $p_i \circ f(a) = p_i(\bigvee_{s\in S} \Downarrow \{q_i(s)\}_{i\in I}) = \bigvee_{s\in S} h_i\{q_i(s)\}_{i\in I} = \bigvee_{s\in S} q_i(s) = q_i(\bigvee_{s\in S} s) = q_i(a)$.

Secondly, f is a **CL**-morphism. Obviously, f preserves finite unions. For any family $\{b^m\}_{m\in M} \subset B$, it suffice to show $\bigvee_{t\in T} f'(t) = \bigvee_{s\in S} f'(s)$, where $T = \{t \mid t \leq \bigvee_{m\in M} b^m$ and $t \in Copr(B)\}$, $S = \{s \mid s \leq b^m$ for some $m \in M$ and $s \in Copr(B)\}$. Clearly, $\bigvee_{t\in T} f'(t) \supset \bigvee_{s\in S} f'(s)$. Let $t \in T$, we will show that $\{q_i(t)\}_{i\in I} \in \bigvee_{s\in S} f'(s)$, which implies $\bigvee_{t\in T} f'(t) = \bigvee_{s\in S} f'(s)$. Assume that $t_0 \in T$ and $\{q_i(t_0)\}_{i\in I} \notin \bigvee_{s\in S} f'(s)$. Then exists a $P = \cup_{d\in D} \Downarrow d \in \mathcal{B}(\{q_i(t_0)_{i\in I}\})$ such that $\cup_{s\in S} f'(s) \subset P$ (and thus $\vee_{s\in S} f'(s) \subset P$), where D is a finite subset of A. We consider the following three cases:

Case 1. $|D| = 0$, where $|D|$ is the cardinality of D. Then $P = \emptyset = S$, and thus $\bigvee_{t\in T} f'(t) = \bigvee_{s\in S} f'(s)$.

Case 2. $|D| = 1$, i.e. $P = \Downarrow d$ for some $d = \{d_i\}_{i\in I} \in A$. $\{q_i(t_0)\}_{i\in I} \notin P$ implies $q_j(t_0) \not\leq d_j$ for some $j \in I$. As $q_j(t_0) \leq q_j(\bigvee T) = q_j(\bigvee S) = \bigvee_{s\in S} q_j(s)$, $\bigvee_{s\in S} q_j(s) \not\leq d_j$. Then $q_j(s_0) \not\leq d_j$ for some $s_0 \in S$. It follows that $f'(s_0) \not\subset P$. This is a contradiction because $\bigcup_{s\in S} f'(s) \subset P$.

Case 3. $|D| \geq 2$. For simplicity, we only consider the case $|D| = 2$, i.e. $P = \Downarrow c \cup \Downarrow d$, where $c = \{c_i\}_{i\in I} \neq d = \{d_i\}_{i\in I}$ and $c, d \in A$. As $\bigvee_{s\in S} f'(s) \subset P$ and $f'(s) \in Copr(\widetilde{A})$ $(\forall s \in S)$, $f'(s) \subset \Downarrow c$ or $f'(s) \subset \Downarrow d$ $(\forall s \in S)$. Let $S_1 = \{s \in S \mid f'(s) \subset \Downarrow c\}, S_2 = \{s \in S \mid f'(s) \subset \Downarrow d\}, x = \bigvee S_1$ and $y = \bigvee S_2$. Then $S = S_1 \cup S_2$ and $\bigvee S = x \vee y$. First, we show $f(x) \subset \Downarrow c$. Suppose that $f(x) \not\subset \Downarrow c$, then $f'(r) = \Downarrow \{q_i(r)\}_{i\in I} \not\subset \Downarrow c$ for some $r \in \downarrow x \cap Copr(B)$ since $Copr(L)$ is a $\vee$-generating set of B. It follows that $q_j(r) \not\leq c_j$ for some $j \in I$, and thus $\bigvee_{s\in S_1} q_j(s) = q_j(x) \not\leq c_j$ and $q_j(s_0) \not\leq c_j$ for some $s_0 \in S_1$. Hence $f'(s_0) \not\subset \Downarrow c$, which contradicts to the definition of S_1. Therefore $f(x) \subset \Downarrow c$. Analogously, $f(y) \subset \Downarrow d$. Then $\bigvee_{t\in T} f'(t) = f(\bigvee T) = f(x \vee y) = f(x) \vee f(y) = f(x) \cup f(y) \subset \Downarrow c \cup \Downarrow d = P$. This contradicts to $\{q_i(t_0)\}_{i\in I} \notin P$.

For every $b \in Copr(B)$, $f(b) = f'(b) = \{q_i(b)\}_{i\in I}$. Then f preserves nonzero co-prime elements, and thus f is a **CL**-morphism.

Uniqueness. Let $h : B \longrightarrow \widetilde{A}$ also be a **CL**-morphism such that $p_i \circ h = q_i$ $(\forall i \in I)$. Then $h(a) = \bigvee_{b\in N} h(b)$, where $N = \{b \in Copr(B) \mid b \leq a\}$. By $p_i \circ h(b) = q_i(b)$ and the definition of p_i, $g_* \circ h(b) = \{q_i(b)\}_{i\in I}$. Then $h(b) \leq g(\{q_i(b)\}_{i\in I}) = f'(b)$, and thus $h(a) \leq f(a)$. On the other hand, we will show $\{q_i(b)\}_{i\in I} \in h(b)$ $(\forall b \in N)$, which implies $f(a) = h(a)$. By the condition $(*)$, there exists a $P = \cup_{d\in D} \Downarrow d \in \mathcal{B}(\{q_i(b)_{i\in I}\})$ such that $\{q_i(b)\}_{i\in I} \notin P$ but $h(b) \subset P$. We consider the following three cases:

Case 1. $|D| = 0$, where $|D|$ is the cardinality of D. Then $P = \emptyset$ and $a = 0$, and thus $h(a) = f(a)$.

Case 2. $|D| = 1$, i.e. $P = \Downarrow d$ for some $d = \{d_i\}_{i\in I} \in A$. Then $q_j(b) \not\leq d_j$ for some $j \in I$. That is $p_j \circ h(b) \not\leq d_j$. By the definition of p_j, $g_* \circ h(b) \not\leq d$. It follows $h(b) \not\subset g(d) = P$. This is a contradiction because $h(b) \subset P$.

Case 3. $|D| \geq 2$. For simplicity, we only consider the case $|D| = 2$, i.e. $P = \Downarrow c \cup \Downarrow d$, where $c = \{c_i\}_{i\in I} \neq d = \{d_i\}_{i\in I}$ and $c,\ d \in B$. Similar to the proof of **Case 2**, we can prove $h(b) \not\subset \Downarrow c$ and $h(b) \not\subset \Downarrow d$. Because of $b \in Copr(B), h(b) \not\subset \Downarrow c \cup \Downarrow d = P$. This is a contradiction because $h(b) \subset P$.

Acknowledgements. This paper is supported by the Foundation of Shaanxi Province (Grant No. 2009JK452) and the Foundation of XPU (Grant No. 09XG10).

References

1. Adámek, J., Herrlich, H., Strecker, G.E.: Abstract and Concrete Categories. John Wiley & Sons, New York (1990)
2. Gierz, G., Hofmann, K.H., Keimel, K., et al.: A Compendium of Continuous Lattices. Springer, Berlin (1980)
3. Wei, H., Ming, L.Y.: Inverse limits in category of locales. Science in China (Ser. A) 41, 476–482 (1998)
4. Wei, H., Plewe, T.: Directed inverse limits of spatial locales. Proceedings of the American Mathematical Society 130, 2811–2814 (2002)
5. Hai-Yang, L., Sheng-Gang, L.: The inverse limits and direct limits in category of CoFrm. Journal of Northeast Normal University (Natural science edition) 38, 11–14 (2006) (in Chinese)
6. Isbell, J.: Direct limits of meet-continous lattices. Journal Pure Applied Algebra 23, 33–35 (1982)
7. Johnstone, P.T.: Stone Spaces. Cambridge University Press, Cambridge (1982)
8. Papert, S.: Which distributive lattices are lattice of closed sets? Mathematical Proceedings of the Cambridge Philosophical Society 55, 172–176 (1959)
9. Sheng-Gang, L., Hai-Yang, L.: Some properties of categories CL and CLPair. Applied Categorical Structures 17, 103–117 (2009)
10. Bin, Z.: Limits in the category of molecular lattices. Acta Math. Sinica 40, 411–418 (1997) (in Chinese)
11. Bin, Z.: The inverse limits in category of topological molecular lattices. Fuzzy Sets and Systems 118, 547–554 (2001)

Relative Density Weights Based Fuzzy C-Means Clustering Algorithms

Jin-hua Chen and Xiao-yun Chen

College of Mathematics and Computer Science,
Fuzhou University, Fuzhou 350108, China
e-mail: c_xiaoyun@21cn.com

Abstract. Fuzzy C-means (FCM) clustering algorithm tries to get the memberships of each sample to each Cluster by optimizing an objective function, and then assign each of the samples to an appropriate class. The Fuzzy C-means algorithm doesn't fit for clusters with different sizes and different densities, and it is sensitive to noise and anomaly. We present two improved fuzzy c-means algorithms, Clusters-Independent Relative Density Weights based Fuzzy C-means (CIRDWFCM) and Clusters-Dependent Relative Density Weights based Fuzzy C-means (CDRDWFCM), according to the various roles of different samples in clustering. Several experiments of them are done on four datasets from UCI and UCR. Experimental results shows that this two presented algorithms can increase the similarity or decrease the iterations to some extent, and get better clustering results and improve the clustering quality.

Keywords: Cluster Analysis, Fuzzy C-means, Fuzzy Pseudo-partition, Relative Density Weights, Cluster Similarity.

1 Introduction

Cluster analysis [1] is the work of assigning a set of unlabeled samples into several different groups (called clusters) in some way, so that samples in the same group are similar as much as possible. At present, cluster analysis has been widely used in many practical problems in real life, such as image processing, information retrieval, data analysis, pattern recognition, data mining etc. Whether the purpose of clustering is understanding or application, cluster analysis is occupying a very important position. Sometimes, cluster analysis is a good beginning of other purposes of analysis. The clustering results can be used in data preprocessing, data classification, anomaly detection. Considering whether we have the demand to assign each sample to a certain class strictly and accurately, cluster analysis can be divided into hard clustering and fuzzy clustering [2]. The traditional methods of cluster analysis are the former, that is, hard clustering and they have an either-or feature. Objective things mostly have some fuzzy nature, for there is not a clear boundary between properties of things and they are not either-or, so cluster analysis methods use the expression of uncertainty, which is

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 459–466.
springerlink.com

more suitable for the nature of objective reality and more objectively reflect the real world. Thus, in-depth study and exploration of the fuzzy clustering has become the main content of cluster analysis researching. Fuzzy C-means (FCM) clustering algorithm wins the most wide application and research in fuzzy clustering.

FCM algorithm, which introduces the fuzzy set theory and fuzzy logic into the K-means, is in fact the fuzzy version of the traditional K-means clustering algorithm. The adopted term, fuzzy pseudo-partition, reflects the degree of each sample belonging to each cluster (called membership). Although the calculating intensity is slightly higher, there are many issues in FCM, which are the same with those existing in K-means [3]. These problems are: A, the clustering results and the convergence rate are greatly influenced by the initial value, which may easily result in a local optimal solution, especially when the number of clusters is bigger; B, when it comes to non-spherical clusters, clusters of different sizes and clusters with greatly difference in density, it shows a low performance; C, it is sensitive to noise and outliers.

To solve the problems B and C pointed out in the above, according to the various roles of different samples in clustering, we add an appropriate weighting factor in the convergence process using the relative density weights. According the weighting factor is dependent or independent on a particular cluster, two relative density weights based fuzzy c-means clustering algorithms are presented, which are Clusters-Independent Relative Density Weights based Fuzzy C-means (CIRDWFCM) and Clusters-Dependent Relative Density Weights based Fuzzy C-means (CDRDWFCM).

2 Fuzzy C-Means (FCM) Algorithms

Let $X=\{X_1, X_2, \cdots, X_m\}$ be the sample set, so that X contains m samples. It assumes that the number of cluster q is known a priori, where q is a positive integer bigger than 1. U is the fuzzy matrix with q rows and m columns, where element U_{ki} represents the membership of the *i*th sample belonging to the *k*th cluster, so U is also called membership matrix. Moreover, p is the fuzziness index.

Let the objective function of FCM O(U,C) be the error sum of squares [3], as follows:

$$O(U,C) = \sum_{i=1}^{m} \sum_{k=1}^{q} U_{ki}{}^{p} d_{ik}{}^{2} \tag{1}$$

where C is a set of q cluster centers, C_k means the *k*th cluster center (or centroid), d_{ik} is the distance between the *i*th sample and the *k*th cluster center. The FCM clustering of sample set X is to minimizes the objective function O(U,C) with the constraint that the sum of membership of each sample X_i $\{i=1, 2, \cdots, m\}$ belonging to all clusters equals 1.

By minimizing the objective function O(U,C) using the Lagrange multiplier method [4], as we adopt the following updating equations used in [3] for U and C.

$$U_{ki} = \frac{1}{\sum_{j=1}^{q} (d_{ik}{}^{2} / d_{ij}{}^{2})^{1/(p-1)}} \tag{2}$$

$$C_k = \sum_{i=1}^{m} U_{ki}{}^{p} X_i / \sum_{i=1}^{m} U_{ki}{}^{p} \tag{3}$$

The detailed steps of FCM are described as follows:

Step 1: Appoint the fuzzy index p, and determine the initial fuzzy pseudo-partition, that is to initialize the membership matrix U;

Step 2: Update all cluster centers C_k {k = 1, 2, ··· , q} according to the fuzzy pseudo-partition U and equation (3), and then recalculate the fuzzy pseudo-partition U using equation (2);

Step 3: Repeat Step 2, until each cluster centroid does not change any more or change of each centroid is less than a given threshold.

In the algorithm steps described in the above, the loop termination condition in Step 3 can be replaced as that the change of the objective function O(U,C) or the change of error is no longer greater than a appointed threshold, or the changes of absolute value of all elements in the membership matrix are all less than a given threshold. For each sample Xi, it has q memberships correspond to q clusters. We finally assign the sample Xi to the cluster membership of which is the greatest.

3 Relative Density Weights Based Fuzzy C-Means Clustering Algorithms

Aiming at solving the mentioned problems of FCM, Jin-Liang Chen et al. propose an improved method in [5]. Considering the various roles of different samples in clustering, they add an appropriate weighting factor in the convergence process of clustering. Their weighting factor depends on the distance between each pair of samples, so it is an invariant. A.H. Hadjahmadi et al. consider changing weights based method in [6] behind the idea in [5]. Their weighting factor is not an invariant, because it is changed as the clusters changed during the convergence process. Both the weighting factor in [5] and [6] take only the distance measure into account, without considering the relative density around each sample, and thus can not commendably reflect the importance of a sample towards each cluster or its importance for clustering. Therefore, considering the relationship of relative density among samples and clusters and according to whether the weighting factor is variable during the clustering process, we propose two improved method of FCM, Clusters-Independent Relative Density Weights based Fuzzy C-means (CIRDWFCM) algorithm and Clusters-Dependent Relative Density Weights based Fuzzy C-means (CDRDWFCM) algorithm.

3.1 Clusters-Independent Relative Density Weights Based Fuzzy C-Means (CIRDWFCM) Clustering Algorithm

We use a fixed value W obtained before clustering process, not during the convergence process, to measure the importance of each sample. W is independent of each cluster, that is, in every iteration, W is never changed. The m weights $\{W_i\}$, i = 1, ···, m corresponding to m samples $\{X_i\}$, i = 1, ···, m in sample set X, make up of a one-dimensional matrix W, which we call weighting factor matrix.

According to W, we modify the objective function O(U,C), like [5], into the following equation:

$$O(U,C) = \sum_{i=1}^{m} \sum_{k=1}^{q} W_i U_{ki}{}^{p} d_{ik}{}^{2} \tag{4}$$

The Lagrange multiplier method [4][5] is used to minimize the objective function O(U,C) and then derive a new optimal solution of U and C. U has the same shape as is shown in equation (2), while C is determined by the updating equation (5) which is used in [5].

$$C_k = \sum_{i=1}^{m} W_i U_{ki}{}^{p} X_i / \sum_{i=1}^{m} W_i U_{ki}{}^{p} \tag{5}$$

Behind ideas based on relative density, we use the relative density around each sample to weigh the importance of a sample for clustering, that is, weight of each sample is determined by its relative density. Firstly, we give three definitions as follows:

Definition 1. t-Neighbor Set of sample x
In sample set X, the t-Neighbor Set of sample x, denoted as N(x, t), is a set of samples meeting the following two conditions:

(1) N(x, t) contains at least t samples;

(2) Let d_max be the maximum distance between samples in N (x, t) and sample x. Then there are at most t-1 samples, distance between each of which and x is less than d_max.

Definition 2. t-Neighbor Density of sample x
In sample set X, let N(x, t) be the t-Neighbor Set of sample x defined by Definition 1. We define the t-Neighbor Density of sample x, which is denoted as $\rho(x,t)$*, as the inverse average distance[3] between all samples in N(x, t) and sample x, that is*

$$\rho(x,t) = \left[\frac{\sum_{y \in N(x,t)} d(x,y)}{|N(x,t)|} \right]^{-1} \tag{6}$$

where N(x,t) is the t-Neighbor Set of sample x, y is one of the samples in N(x,t), d(x,y) is the distance between sample x and sample y, and |N(x,t)| is the number of samples in N(x,t).

Definition 3. t-Neighbor Relative Density of sample x
In sample set X, the t-Neighbor Relative Density of sample x, denoted as $\gamma\_\rho(x,t)$ *, is defined by the following formula based on Definition 1 and Definition 2:*

$$\gamma\_\rho(x,t) = \frac{\rho(x,t)}{\sum_{y \in N(x,t)} \rho(y,t) / |N(x,t)|} \tag{7}$$

where N(x,t) is the t-Neighbor Set of sample x, y is one of the samples in N(x,t), |N(x,t)| is the number of samples in N(x,t) and $\rho(x,t)$ *and* $\rho(y,t)$ *are respectively the t-Neighbor Density of sample x and y.*

Based on the definition above, we can determine each element W_i *in weighting factor matrix W as follows:*

$$W_i = \gamma\_\rho(X_i, t) \tag{8}$$

where X_i *is the ith sample of sample set X, t is a control parameter, which shows that we adopt t-Neighbor Relative Density.*

3.2 Clusters-Dependent Relative Density Weights Based Fuzzy C-Means (CIRDWFCM) Clustering Algorithm

In CDRDWFCM, we also add a weighting factor W to the objective function. Unlike the CIRDWFCM, here the weighting factor matrix W is not an invariant and is not derived before the clustering process, but is variable during the clustering process. W is a q rows and m columns matrix, where each element is related to a specific cluster and element W_{ki} represents the influence factor of the *i*th sample to the *k*th cluster.

Adding this type of weights, the CDRDWFCM algorithm aims to minimize the following objective function according to [6]:

$$O(U,C) = \sum_{i=1}^{m} \sum_{k=1}^{q} W_{ki} U_{ki}{}^{p} d_{ik}{}^{2} \tag{9}$$

Due to the dependence of weighting factor on specific clusters and each weight is no more independent of each cluster center, to minimize O (U, C) becomes very complex. For simplicity of calculation, we approximatively use the assumption that the partial derivative of weight to cluster centroid equals 0. Then, based on this assumption, using the Lagrange multiplier method [4], we adopt the updating equations used in [6] for U and C:

$$U_{ki} = \frac{1}{\sum_{j=1}^{q} (W_{ki} d_{ik}{}^2 / W_{ji} d_{ij}{}^2)^{1/(p-1)}} \tag{10}$$

$$C_k = \sum_{i=1}^{m} W_{ki} U_{ki}{}^p X_i / \sum_{i=1}^{m} W_{ki} U_{ki}{}^p \tag{11}$$

Considering the idea of relative density again, we give a new weight W which is not only related to the sample itself, but also related to the specific cluster. Each element of W, W_{ki}, is computed as follows:

$$W_{ki} = \frac{\min\{\rho(X_i, t), \rho\_avg(C_k, t)\}}{\max\{\rho(X_i, t), \rho\_avg(C_k, t)\}} \tag{12}$$

where C_k is the *k*th cluster center (or centroid), t is a control parameter, which shows that we adopt t-Neighbor Relative Density, $\rho(X_i, t)$ and $\rho(C_k, t)$ are respectively the t-Neighbor Density of sample X_i and C_k, and $\rho\_avg(C_k, t)$ equals the average t-Neighbor Density of all samples in t-Neighbor Set of sample C_k.

Weighting factor defined in equation (12) considers both the local density around each sample itself and the local density around each cluster center, and then determine the importance (weight) of one sample to one cluster by the two densities.

4 Experimental Results and Analysis

We analyze the ability of CIRDWFCM and CDRDWFCM by comparison with other related algorithms. This experiment adopts two real data sets from UCI Machine Learning Repository [7][5] and two real data sets from UCR Time Series Data Mining Archive [8].

4.1 Experimental Data Descriptions

The four data sets adopted in this experiment are Iris and sonar from UCI, and Lighting2 and Gun_point from UCR. They are all real data sets, and the last two are time series data sets, as is shown in Table 1.

Table 1. Experimental Data Descriptions

Data sets	Iris	sonar	Lighting2	Gun_point
Samples Number	150	208	60	50
Attributes Number	4	60		
Series Length			637	150
Classes Number	3	3	2	2
Distribution	50+50+50	97+111	20+40	24+26

4.2 Clustering Results Evaluation

In [9], T.W. Liao proposes a method to evaluate the clustering results, which adopt the index called cluster similarity measure. The greater the cluster similarity measure is, the better the clustering quality is. Let k and q respectively be the real number of classes and the number of clusters in clustering results. Let G and C be the set of k ground truth clusters and the set of q clusters obtained by a clustering method respectively. The similarity between the *s*th real cluster and the *t*th obtained cluster is defined as

$$Similarity(G_s, C_t) = 2 \times \frac{|G_s \cap C_t|}{|G_s| + |C_t|} \tag{13}$$ [9]

where $|\cdot|$ denotes the cardinality of the elements in the set.

The final cluster similarity measure of a clustering result is defined as equation (14).

$$Similarity(G, C) = \frac{1}{q} \sum_{s=1}^{q} \max_{1 \le t \le q} Similarity(G_s, C_t) \tag{14}$$ [9]

4.3 Experimental Results and Analysis

In order to analyze the clustering ability of the proposed algorithms, we compare them with the FCM and DWFCM algorithms. The experiment need to set a parameter t, here we assume it as 1/3 of the average number of samples for all clusters. For comparison, we adopt the cluster similarity measure as the clustering results evaluation index, do the experiment repeatedly for 10 times and then fetch the best results as shown in Table 2.

Table 2. Comparison of Clustering Results

Data sets		Iris	sonar	Lighting2	Gun_point
FCM	Iterations Number	17	51	71	10
	Cluster Similarity	0.886109	0.552792	0.534349	0.557166
DWFCM	Iterations Number	12	47	73	9
	Cluster Similarity	0.885741	0.552792	0.534349	0.557166
CIRDWFCM	Iterations Number	11	19	16	7
	Cluster Similarity	0.892948	0.552874	0.534349	0.557166
CDRDWFCM	Iterations Number	15	96	46	9
	Cluster Similarity	0.898775	0.557651	0.558035	0.575758

From Table 2, we see: (1) For data sets Iris and sonar, CIRDWFCM and CDRDWFCM cause the greater cluster similarity compared to FCM and DWFCM; (2) For the other two data sets, CDRDWFCM still has a cluster similarity greater than that FCM and DWFCM have, and the number of iterations CIRDWFCM needed is obviously less than FCM and DWFCM, although CIRDWFCM doesn't produce a higher cluster similarity. Therefore, the proposed algorithms can get greater cluster similarity or reduce the number of iterations, that's to say, they can improve the clustering performance. Compared to FCM and DWFCM, the proposed algorithms show better clustering performance.

5 Conclusions

The traditional Fuzzy C-means (FCM) clustering algorithm has some problems. Aiming at improving it, in this paper, we propose two algorithms CIRDWFCM and CDRDWFCM based on the idea of weighting factor and the concept of relative density. The experimental results show that the proposed methods have the stronger clustering ability.

Acknowledgements. This research was supported by the National Natural Science Foundation of China under Grant No. 60573076 and the Program for New Century Excellent Talent in Fujian Province (XSJRC2007-11).

References

1. Han, J., Kamber, M.: Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers, San Francisco (2001)
2. Zhang, M., Yu, J.: Fuzzy partitional clustering algorithms. Journal of Software 15(6), 858–869 (2004)
3. Tan, P.-N., Steinbach, M., Kumar, V.: Introduction to Data Mining. Addison Wesley, New York (2006)
4. Cheng, W.X., Min, S.: The basic principles of finite element method and numerical method. Tsinghua University Press, Beijing (1997)
5. Chen, J.L., Wang, J.H.: A new robust clustering algorithm—density-weighted fuzzy c-means. In: Proceedings of the Conference IEEE SMC, vol. 3, pp. 90–94 (1999)
6. Hadjahmadi, A.H., Homayounpour, M.M., Ahadi, S.M.: Robust Weighted Fuzzy C-Means Clustering. In: IEEE International Conference on Fuzzy Systems, pp. 305–311 (2008)
7. Newman, D.J., Hettich, S., Blake, C.L., Merz, C.J.: UCI Repository of machine learning databases (1998), `http://www.ics.uci.edu/~mlearn/MLRepository.html`
8. Keogh, E., Xi, X., Wei, L., Ratanamahatana, C.A.: The UCR Time Series Classification / Clustering (2006), `http://www.cs.ucr.edu/~eamonn/time_series_data/`
9. Liao, T.W.: Clustering of time series data—a survey. Pattern Recognition 38, 1857–1874 (2005)

A Generalization of the Lowen Functor ω_L

Ling-xia Lu[1] and Wei Yao[2]

[1] School of Mathematics and Science, Shijiazhuang University of Economics, Shijiazhuang 050031, China
e-mail: lulingxia1214@gmail.com

[2] Department of Mathematics, Hebei University of Science and Technology, Shijiazhuang 050018, China
e-mail: yaowei0516@163.com

Abstract. This paper generalize the Lowen functor based on a complete lattice with an approximating relation (with the property of interpolation). It is shown that, on any completely distributive lattice, the Lowen functor ω_L can not only defined by $\not\leq$ relation, but also by the way below relation $\ll$ and the wedge below relation $\lhd$.

Keywords: Lowen functor; completely distributive lattice; way-below relation; wedge-below relation.

1 Introduction

The Lowen functors (ω, ι) were introduced by Lowen [8] in order to study the relations between the category of topological spaces **Top** and the category of fuzzy topological spaces [0,1]-**Top**. Later, several authors extended this adjunction to completely distributive lattices [7, 9] and even to complete lattices [4, 5].

Let (X, T) be a topological space and L a completely distributive complete lattice. $A \in L^X$ is called a lower semicontinuous function w.r.t. T if for any $a \in L$, $\iota_a(A) = \{x \in X |\ A(x) \not\leq a\} \in T$. The family of all lower semicontinuous functions w.r.t. T is denoted by $\omega_L(T)$, which forms a stratified L-topology on X. It's easy to see that $A \in L^X$ is lower semicontinuous iff $A : (X, T) \longrightarrow (L, \nu(L))$ is continuous, where $\nu(L)$, generated by the subbasis $\{L\backslash \downarrow a |\ a \in L\}$, is called the upper topology on L.

Lowen's original definition is the case $L = [0, 1]$. Now, let's focus on the lattice [0,1]. Obviously, [0,1] is a completely distributive lattice and thus is also a continuous lattice. The binary relation "less than $<$" is exactly the way below relation and the wedge below relation on [0,1], both of which are approximating relations (with the property of interpolation). It's natural to ask that,

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 467–471.
springerlink.com

For a completely distributive lattice L, can we generalize the Lowen functor by means of way below relation or wedge relation? or *Can the Lowen functors be characterized by the wedge relation and/or by the way below relation?*

The aim of this paper is to generalize Lowen functors based on a complete lattice with an approximating relation (with the property of interpolation) and then study the special cases for the way-blow and wedge-below relations.

2 Preliminaries

In this paper, a lattice is always assumed to be complete.

The set of all L-topologies on a set X is a complete lattice under partial order of inclusion. For all $\mathcal{S} \subseteq L^X$, $\langle\langle\mathcal{S}\rangle\rangle$ denotes the least L-topology that contains $\mathcal{S}$, called the L-topology generated by $\mathcal{S}$.

An element $a \in L$ is called co-prime if for any $b, c \in L$, $a \leq b \vee c$ implies $a \leq b$ or $a \leq c$. The set of all non-zero co-prime elements of L is denoted by $J(L)$.

For any $a, b \in L$, a is said to be way below (resp., wedge below) b, in symbols $a \ll b$ (resp., $a \lhd b$), if for any directed subset (resp., any subset) $D \subseteq L$, $b \leq \bigvee D$ always implies $a \leq d$ for some $d \in D$. Put $\Uparrow a = \{x \in L \mid x \ll a\}$ (resp., $\beta(a) = \{x \in L \mid x \lhd a\}$) for each $a \in L$. L is called continuous (resp., completely distributive) iff for any $a \in L$, $a = \bigvee \Uparrow a$ (resp., $x = \bigvee \beta(a)$). The Scott topology $\sigma(L)$ on a continuous lattice L is the topology generated by the basis $\{\Uparrow a \mid a \in L\}$, the corresponding topological space is denoted by $\Sigma(L)$.

The completely distributive law is described as:

(CD) $\quad \bigwedge\limits_{j \in J} \bigvee A_j = \bigvee \{ \bigwedge\limits_{j \in J} \varphi(j) \mid \varphi \in \prod\limits_{j \in J} A_j \}$

for arbitrary $A_j \subseteq L$ and arbitrary index set J.

A complete lattice L is a completely distributive iff it satisfies the completely distributive law.

Lemma 1. *Let L be a completely distributive lattice. Then*
(1) for any $a \in L$, $a = \bigvee \beta(a) = \beta^*(a)$, where $\beta^*(a) = \beta(a) \cap J(L)$;
(2) for any $a \in J(L)$, $b \in L$, $a \ll b$ iff $a \lhd b$;
(3) $\nu(L) = \sigma(L)$.

L is called a frame if it satisfies the following infinitely distributive law:

(ID) $\quad a \wedge (\bigvee B) = \bigvee\limits_{b \in B} (a \wedge b)$

for any $a \in L$ and $B \subseteq L$.

A continuous frame L is a continuous lattice as well as a frame. It's easy to see that a completely distributive lattice is a continuous frame.

3 $\prec$-Lower Semicontinuous Functions

Suppose that $\prec$ a binary relation on L which is less than $\leq$, i.e., for any $a, b \in L,\ a \prec b$ always implies $a \leq b$. For any $A \in L^X$ and any $a \in L$, define $\iota_a^{\prec}(A) = \{x \in X |\ a \prec A(x)\}$.

Theorem 1. *For any $A \in L^X$, $A = \bigvee_{a \in L} \overline{a} \wedge \chi_{\iota_a^{\prec}(A)}$.*

Obviously, $A \geq \bigvee_{a \in L} \overline{a} \wedge \chi_{\iota_a^{\prec}(A)}$. Conversely, for any $x \in X$ and $a \prec A(x)$, we have $\overline{a} \wedge \chi_{\iota_a^{\prec}(A)}(x) \geq a$.

Let (X, T) be a topological space. $A \in L^X$ is called a $\prec$-lower semicontinuous function w.r.t. (X, T) iff for any $a \in L,\ \iota_a^{\prec}(A) \in T$. The family of all $\prec$-lower semi-continuous functions w.r.t. (X, T) is denoted by $\omega_L^{\prec}(T)$. Obviously, on [0,1], a lower semi-continuous function is a $<$-lower semicontinuous function.

Theorem 2. *Each $\prec$-lower semicontinuous function is a lower semicontinuous function.*

Trivial since $\prec$ is approximating, $\prec \subseteq \leq$ and $\iota_a(A) = \bigcup_{b \not\leq a} \iota_b^{\prec}(A)$.

Theorem 3. *(1) For any $a \in L,\ U \in T,\ \overline{a} \wedge \chi_U \subseteq \omega_L^{\prec}(T)$.*
(2) $\omega_L^{\prec}(T) \subseteq \langle\langle\{\overline{a} \wedge \chi_U |\ a \in L,\ U \in T\}\rangle\rangle$.

(1) Let $a \in L,\ A \in T$. For any $b \in L$, if $b \prec a$ then $\iota_b^{\prec}(\overline{a} \wedge \chi_A) = A \in T$ and if $b \not\prec a$ then $\iota_b^{\prec}(\overline{a} \wedge \chi_A) = \emptyset \in T$.

(2) Let $A \in L^X$ be a $\prec$-lower semicontinuous function w.r.t. (X, T). Then $\forall a \in L,\ \iota_a^{\prec}(A) \in T$ and by Lemma 1, $A = \bigvee_{a \in L} \overline{a} \wedge \chi_{\iota_a^{\prec}(A)}$.

Corollary 1. *$\{\overline{a} \wedge \chi_A |\ a \in L,\ A \in T\}$ is a subbase of $\langle\langle\omega_L^{\prec}(T)\rangle\rangle$. If L is a frame, then $\{\overline{a} \wedge \chi_A |\ a \in L,\ A \in T\}$ is a base. Thus $\langle\langle\omega_L^{\prec}(T)\rangle\rangle$ is a fixed (stratified) L-topology on X.*

Theorem 4. *$\omega_L^{\prec}$: **Top** $\longrightarrow$ **S**L**-Top**, $(X, T) \mapsto (X, \langle\langle\omega_L^{\prec}(T)\rangle\rangle)$, is a functor*

For any $a \in L$, let $\uparrow_{\prec} a = \{x \in L |\ x \prec a\}$ and $\sigma_{\prec}(L)$ denote the (crisp) topology on X generated by the subbasis $\{\uparrow_{\prec} a |\ a \in L\}$.

Theorem 5. *$A \in \omega_L^{\prec}(T)$ iff $A : (X, T) \longrightarrow (L, \sigma_{\prec}(L))$ is continuous.*

Trivial since $A^{-1}(\uparrow_{\prec} a) = \iota_a^{\prec}(A)$.

Corollary 2. *If L is a continuous lattice, then $A \in \omega_L^{\ll}(T)$ iff $A : (X, T) \longrightarrow (L, \sigma(L))$ is continuous.*

Theorem 6. *(Warner [11]) For a continuous frame L, the family of all continuous functions from (X, T) to $(L, \sigma(L))$ is an L-topology on X.*

Corollary 3. *Let L be a continuous frame. Then $\omega_L^{\ll}(T)$ already forms the L-topology on X.*

Corollary 4. *If L is completely distributive, then $A \in L^X$ is lower semicontinuous iff A is $\ll$-lower semicontinuous, i.e., $\omega_L = \omega_L^{\ll}$.*

As a corollary of Theorem 2, we have.

Corollary 5. *If L is completely distributive, then $\omega_L^{\lhd}(T) \subseteq \omega_L(T)$.*

Lemma 2. *[1] $\omega_L^{\ll}(T) \subseteq \omega_L^{\lhd}(T)$.*

We only need to show that $\iota_a^{\lhd}(A) = \bigcup\limits_{b \in \uparrow a \cap J(L)} \iota_b^{\ll}(A)$. For any $x \in \iota_a^{\lhd}(A)$, we have $a \lhd A(x)$ and then $a \le b$ for some $b \in \beta^*(A(x))$. Then $b \in \uparrow a \cap J(L)$ and $b \ll A(x),\ x \in \iota_b^{\ll}(A)$. Conversely, for any $b \in \uparrow a \cap J(L)$ and any $x \in \iota_b^{\ll}(A)$, we have $a \le b \lhd A(x)$ since $b \in J(L)$. Hence $a \lhd A(x)$ and $x \in \iota_a^{\lhd}(A)$.

Corollary 6. $\omega_L^{\lhd} = \omega_L^{\ll} = \omega_L$.

4 $\omega_L = \omega_L^{\ll}$ for Distributive Continuous Lattices

A complete lattice L is called a continuous frame if L is both a frame and a continuous lattice.

In this section, L always denotes a completely distributive lattice.

Theorem 7. $\{\Uparrow a |\ a \in M(L)\}$ *forms a subbasis of $\sigma(L)$, where $\Uparrow = \uparrow_{\ll}$.*

For any $U \in \sigma(L)$ and any $u \in U$, there exists $v \in U$ such that $v \ll u$ since every completely distributive lattice is a continuous lattice. Since $v = \bigvee\{x \in M(L) |\ x \le v\} \in U$, there exits $x_1, \cdots, x_n \in M(L) \cap \downarrow v$ such that $x_1 \vee \cdots \vee x_n \in U$. It's easy to show that $u \in \Uparrow x_1 \cap \cdots \cap \Uparrow x_n = \Uparrow (x_1 \vee \cdots \vee x_n) \subseteq U$. Hence $\{\Uparrow a |\ a \in M(L)\}$.

Theorem 8. *[11] $A \in \omega_L(T)$ iff $A : (X, T) \longrightarrow (L, \sigma(L))$ is continuous.*

$\Longrightarrow$ For any $a \in M(L)$, we only need to show that $A^{-1}(\Uparrow a) = \iota_{\bigvee L \setminus \uparrow a}(A)$. In fact, for any $x \in A^{-1}(\Uparrow a)$, we have $a \ll A(x)$. If $A(x) \le \bigvee L \setminus \uparrow a$, then $a \le d$ for some $d \in L \setminus \uparrow a$ ($L \setminus \uparrow a$ is an upper directed set), which is a contradiction. Thus $A(x) \not\le \bigvee L \setminus \uparrow a$ and $x \in \iota_{\bigvee L \setminus \uparrow a}(A)$. Conversely, suppose that $x \in \iota_{\bigvee L \setminus \uparrow a}(A)$, then $A(x) \not\le \bigvee L \setminus \uparrow a$. Let D be a upper directed set and $A(x) \le \bigvee D$, then $\bigvee D \not\le \bigvee L \setminus \uparrow a$ and there exist $d \in D$ such that $d \not\le \bigvee L \setminus \uparrow a$. It follows that $a \le d$. Thus $a \ll A(x)$ and $x \in A^{-1}(\Uparrow a)$.

$\Longleftarrow$ For any $a \in L,\ \iota_a(A) = A^{-1}(L \setminus \downarrow a) \in T$.

Acknowledgements. This work is supported by the National Natural Science Foundation of China (10926055), the Foundation of Hebei Province (A2010000826, 09276158) the Foundation of HUBUST (XL200821, QD200957) and the Foundation of SJZUE (XN201003). The authors are grateful to Prof. Kubiak for sending them his papers and for his suggestions during preparing this paper.

References

1. Adámek, J., Herrlich, H., Strecker, G.E.: Abstract and Concrete Categories. Wiley, New York (1990)
2. Gierz, G., et al.: Continuous Lattices and Domains. Cambridge University Press, Cambridge (2003)
3. Kotzé, W., Kubiak, T.: Fuzzy topologies of Scott continuous functions and their relation to the hypergraph functor. Quaest. Math. 15, 175–187 (1992)
4. Kubiak, T.: The topological modification of the L-fuzzy unit interval. In: Rodabaugh, S.E., Klement, E.P., Höhle, U. (eds.) Applications of Category Theory to Fuzzy Subsets. Kluwer Academic Publishers, Dordrecht (1992)
5. Kubiak, T.: On L-Tychonoff spaces. Fuzzy Sets and Systems 73, 25–53 (1995)
6. Kubiak, T., de Prada Vicente, M.A.: L-regular topological spaces and their topological modifications. Int. J. Math. Sci. 23, 687–695 (2000)
7. Liu, Y.M., Luo, M.K.: Induced spaces and fuzzy Stone-Čech compactifications. Sinica Ser. A 40, 1034–1044 (1987)
8. Lowen, R.: Fuzzy topological spaces and fuzzy compactness. J. Math. Anal. Appl. 56, 623–633 (1976)
9. Wang, G.J.: Lattice valued lower semicontinuous mappings. J. Shaanxi Normal Univ. 4, 1–7 (1988)
10. Wang, G.J.: Theory of topological molecular lattices. Fuzzy Sets and Systems 47, 351–376 (1992)
11. Warner, M.W.: Fuzzy topology with respect to continuous lattices. Fuzzy Sets and Systems 35, 85–91 (1990)

Remote Sensing Image Classification Based on Fuzzy Entropy Triple I Algorithm

Jun Yan[1], Shui-li Chen[1,2], and Yun-dong Wu[1,2]

[1] School of Science, Jimei University, Xiamen, 361021, China
[2] Research Centre of Image Information Engineering, Jimei University, Xiamen, 361021, China
e-mail: yanjun5201@163.com

Abstract. By analyzing the principle of the supervised classification algorithms, traditional supervised algorithms are either much single for its collected sample, which makes various features from different classes can't be fully characterized, or much larger for its collected sample block, which makes little practical sense and not very good for the results of classification, although including kinds of feature information in the same class. Combining with algorithm of fuzzy entropy triple I, a new classification algorithm of fuzzy entropy triple I based on multi-sample collections is proposed. It is used in ground objects classification and water area extraction of remote sensing image. The results of experiments show that the new algorithm has higher precision, lower false accept rate and false reject rate, and stronger applicability.

Keywords: Remote sensing image, minimum distance discrimination method, Bayes algorithm, fuzzy entropy triple I algorithm.

1 Introduction

With the rapid development of remote sensing technology, remote sensing image has also been widely used. For example, detailed investigation in use on country's land, cover type identification of land, and classifications are all around the object to the use of remote sensing images. Currently, there are two main types of classification algorithms: supervised classification and unsupervised classification. Supervised algorithm includes minimum distance criterion, Fisher linear discrimination, Bayes algorithm, k-nearest neighbor discrimination law etc [1]. However, in these algorithms, collected sample is either single, which does not contain all the characteristics information of the class, or very large block, which makes little significance in the practical application. If the color of heterogeneous is similar, the traditional supervised criterion can not distinguish them. All in all, the results of classification in traditional supervised algorithms are faulty. These problems are due to deficiencies in the algorithm themselves, it either brings a lot of wrong numeracy

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 473–482.
springerlink.com

or lends severe rejection phenomenon, which makes the classification results are not very good.

Based on the truth that human eye is color sensitive and well capable of distinguishing, and remote sensing image includes rich feature information. Aiming at the flaws of traditional supervised algorithms, a new criterion of fuzzy entropy triple I algorithm based on multi-sample collections is proposed and applied into ground objects classification and water area extraction. Compared to the minimum distance discrimination method and the traditional Bayes criterion, this algorithm can improve the classification performance and accuracy.

2 Minimum Distance Discrimination Method

Minimum distance classification is a simple supervised method, which has known the determinate location parameters of surface features in the spectrum space, and then the element is fallen into the model which has the minimum distance by defining it to the center of each pattern class. Based on different definitions of distance, we can get a variety of specific methods. Commonly, many methods are based on Euclidean distance, Mahalanobis distance and so on [1]; The common points can be stated as: assuming the data has m bands, and n classes are described by standard samples $w_1, w_2, \cdots w_n$, according to the principle of minimum distance classification, the distance of the identified element x to the ith class can be defined as follows:

$$D_i(x) = \sqrt{x(x-u_i)^T (x-u_i)} \qquad (i = 1, 2, ..., n) \ , \tag{1}$$

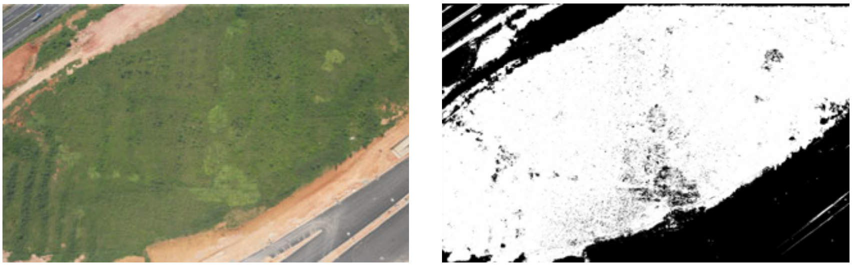
$$D(x) = \sqrt{(x-u)\ \Sigma\ (x-u)} \qquad (i = 1, 2, \cdots, n) \ , \tag{2}$$

(Eq.(1) and (2) are Euclidean distance, Mahalanobis distance, separately[2]), u_i is the mean vector of the ith class, and Σ_i is the covariance matrix of overall distribution. Classification criterion is as following:

$$D(p) < D(p)\ , \quad \forall i \neq j \Rightarrow p \in w\ .$$

From the above, we can see that the calculation of classifier based on the minimum Euclidean distance is very simple. But in general, the accuracy of classification depends on the number of classes and the dispersion of various types. When the models are much more and some kind of spectral spreading is large, the results are not ideal. Although classifications based on the minimum Mahalanobis distance, taking into account the characteristics of the distribution sample, all kinds of general covariance matrix are often difficult to be prevised.

Figure 1(b) is the assorted result of figure 1.(a) based on minimum Euclidean distance. Apparently, there are many scattered grasses in figure 1(b), which should be classified in the class of grass. And in order to prevent the trivial in an unit class and overcome the uncertainty of general covariance matrix in classification based on minimum Mahalanobis distance, the Bayes discrimination [3] is introduced into



(a) (b)

Fig. 1. (a) Original RGB image; (b) the result of grass based on Euclidean distance

the RGB image classification. Taking advantage of the sample's variance matrix instead of the one of the corresponding class makes a better overall effect.

3 Traditional Bayes Criterion

The basic idea of Bayes discrimination method is [1]: the prior probability of each model class should be predetermined before sampling. Then make an amendment to prior knowledge by using the collected samples to get a posterior probability. Finally, assort the unknown samples based on posterior probability. Discriminant analysis on the ideas of Bayes classification, we get Bayes criterion.

If Bayes discrimination is used for image classification, first of all, the prior probability $p(w)$ of model w_i should be determined by experienced experts. Then amend the prior probability by using selected samples, and calculate the class-conditional probability density function $p(x|w_i)$ of unknown element x. For a two types of pattern classification problem, Bayes criterion based on minimum error rate as follows: for element x to be identified, if the probability belongs to pattern class w_1 is greater than it belongs to w_2, we consider that x belongs to w_1, otherwise x belongs to w_2. That is, if $p(w|x) > p(w|x)$, $x \in w_1$, on the contrary, $x \in w_2$.

By the knowledge of probability: $p(x, y) = p(x|y)p(y)$, we know the Bayesian formula:

$$p(w_i|x) = \frac{p(x|w_i)p(w_i)}{p(x)} . \tag{3}$$

Among Eq.(3), $p(x|w_i)$ is the class conditional probability density of x under class w_i. In the Bayes algorithm, we get different results through selecting distinct probability density function. Generally, the Gaussian probability density functions (PDF) is the most interested. The n-dimensional Gaussian density function (PDF) calculated as follows:

$$p(x \mid w_j) = \frac{1}{(2\pi)^{n/2} \mid C_j \mid^{1/2}} e^{-\frac{1}{2}[(x-m_j)^T C_j^{-1}(x-m_j)]}, \tag{4}$$

where, C_j and m_j are covariance matrix and mean vector separately of the family class of models $w_j\ (j = 1, 2, \cdots, c)$, $\mid C_j \mid$ is the determinant of C_j . m_j and C_j are obtained by the followings:

$$m_j = \frac{1}{N_j}\sum_{k=1}^{N_j} w_{jk} \quad , \quad C_j = \frac{1}{N_j - 1}\sum_{k=1}^{N_j}(w_{jk} - m_j)(w_{jk} - m_j)^T , \tag{5}$$

N_j represents the number of sample corresponding to the model class w_j ,and w_{jk} is the kth sample of pattern class w_j . Thus, the criterion of traditional Bayes discrimination method can also be written as:

If $p(x \mid w_1) p(w_1) > p(x \mid w_2) p(w_2)$, then $x \in w_1$; otherwise, $x \in w_2$.

However, Bayes discrimination chooses a single sample also, there is no good to show similar features which exist differences in this situation. And coupling with the subjective determination of the pattern's prior probability, the result of Bayes criterion in the overall classification is better, but there still has some unnecessary defects and contains a lot of wrong identification number. In order to overcome the shortcomings of traditional supervised algorithm and improve the accuracy, this paper does a further improvement based on the Bayes discrimination and proposes fuzzy entropy triple I algorithm.

4 Fuzzy Entropy Triple I Algorithm

4.1 The Basic Idea of Fuzzy Reasoning

In classical propositional operations, we assume that A and B are any two propositions (or formulas), the expression "*If A Then B*" can be written as $A \to B$, then using MP rule, and in the case of knowing A and $A \to B$, we can get B . This reasoning process can be written in the following form:

$$\begin{array}{ll} \text{known} & A \to B \\ \text{and given} & \underline{A\quad} \\ \text{receive} & B. \end{array} \tag{6}$$

But in Eq.(6), where the second line of A is different the one which is in the first row of the containing type " $A \to B$ ",that is, if we replace the second line of A with A^* ,we will get the following inference:

$$\begin{array}{ll} \text{known} & A \to B \\ \text{and given} & \underline{A^*\quad} \\ \text{receive} & B^*, \end{array} \tag{7}$$

where, A is different from A^*. From the point of view of classical logic, Eq.(7) is a pathetic question and can't be answered, because A^* is not A, and all of A, B, A^* are pure forms of symbols. In the case of giving practical meanings to A, B, A^* and thus can be considered the operations of A, B, A^* as well as whether A^* and A are similar or not, it is possible to give the solution to B^* [4]. That's fuzzy reasoning to solve problems. In 1973, L.A.Zadeh proposed CRI algorithm to solve these problems, the form of the computation is:

$$B^*(y) = \sup_{x\in X}[A^*(x) \wedge R(x, y)]$$

$$= \sup_{x\in X}[A^*(x) \wedge R_Z(A(x), B(y))] .$$

Promptly, $B^*(y) = \sup_{x\in X}\{A^*(x) \wedge [A'(x) \vee (A(x) \wedge B(y))]\}$, $y \in Y$, where A, A^* and B, B^* are non-empty sets of fuzzy sets X, Y separately. And $R_Z(A(x), B(y))$ is Zadeh implication operator.

Fuzzy reasoning can also have a more general form:

$$\begin{array}{lll} \text{known} & A_1 \to C_1 & \\ & \ldots\ldots & \\ & A_n \to C_n & \\ \text{and given} & \underline{A^* \quad\quad} & (8) \\ \text{receive} & C^*, & \end{array}$$

where n is the number of inference rules, and we can use the CRI approach to solve B^* also, specific process please see [4].

4.2 Fuzzy Entropy Triple I Algorithm

In many forms of fuzzy inference, we consider the most basic form of reasoning FMP (promptly Eq.(7)), and assume sets $A, A^* \in F(X)$, $B, B^* \in F(Y)$, that is the propositions are represented as Fuzzy Sets. CRI method uses only once conversion in reasoning process, it doesn't consider there would have relationships between $A^* \to B^*$ and $A \to B$ while giving A^* to seek B^*, but simply letting A^* and R complex (compositional rule) to obtain B^*. Through analyzing the characteristics of CRI algorithm, professor Wang proposed the triple I fuzzy algorithm [4]:

Set X and Y are two non-empty sets, and knowing $A, A^* \in F(X)$, $B, B^* \in F(Y)$, then B^*, which makes the next calculation obtain maximum value for all $x \in X$ and $y \in Y$, is the smallest fuzzy set of $F(Y)$

$$(A(x) \to B(y)) \to (A^*(x) \to B^*(y)) . \quad (9)$$

On this basis, [5] uses the fuzzy entropy[6] to measure the fuzzy degree of the results of fuzzy reasoning. For solving the problem likes FMP, a new algorithm of fuzzy reasoning--fuzzy entropy triple I algorithm, is proposed:

Assume X and Y are two non-empty collections, and giving $A, A^* \in F(X)$, $B, B^* \in F(Y)$, then the output B^* of Eq.(7) which makes Eq.(9) obtain maximum value , is the fuzzy set of which has greatest fuzzy entropy. [7] introduced a variety of fuzzy implication operators on fuzzy entropy three-I algorithm, and gives a detailed solving formula. Through remote sensing data as well as traditional supervised algorithm, this paper combines Bayesian probability with minimum distance to establish fuzzy inference rules. Implication $R(a,b) = (1-a+b) \wedge 1$ is adopted, and the solution of the fuzzy entropy triple I described as [7]:

$$B^*(y) = \sup_{x \in E_y}\{R_0(A(x), B(y)) + A^*(x) - 1\} \vee \frac{1}{2}, \qquad y \in Y \tag{10}$$

and

$$E_y = \{x \in X \mid R(A(x), B(y)) + A^*(x) > 1\}. \tag{11}$$

At target-pattern classification, we can joint the multi sample blocks' data in similar class to pursue higher accuracy in results based on fuzzy reasoning which can be amended n inference rules. In real terrain classification, at first, we can view the remote sensing image as a fuzzy set [8], and calculate the mean vector M_{ij} and covariance matrix C_{ij} of the collected sample (which can be seen as fuzzy subsets of the whole image) in model class using Eq.(5), where i means the ith model-class and j expresses the jth sample-plate of the ith model-class. And then define the distance $D_{ij}(x)$ which obtained by the gray-scale vector of element x and the sample mean vector (both have to be normalized) as a membership of x corresponds to the pattern class [2]. Similarly, $P_{ij}(x)$ is easy calculated by Eq.(4), and it can be seen as a membership also. Thus, the use of information of the elements u in the sample block, we can calculate $D_{ij}(u)$ and $P_{ij}(u)$. Now, it determines a fuzzy inference rule: if the distance of a certain element u to M_{ij} is $D_{ij}(u)$, then the probability of which attaches to the corresponding model-class is $P_{ij}(u)$. Hence, if the gray-scale vector of be identified element x has given, we can get $P_{ij}(x)$, and then use the fuzzy entropy triple I algorithm and Eq.(10) to solve $P_{ij}^*(x)$. Finally, combine the obtained $P_{ij}^*(x)$ $(j = 1,2,3,\cdots m_i)$, where m_i is the samples number of the ith corresponding model class) of all the sample blocks of similar class, and obtain the probability $P_i^*(x)$ of the identifying element x attach to each pattern class $(i = 1,2,3\cdots N)$, N is the number of pattern class) by using Kleene-Dienes implication operator. At last, classify the target patterns according to the following criterion:

If $P = \max\{P_1^*(x), P_2^*(x), \cdots, P_N^*(x)\} = P_k^*(x)$, x will be determined to belong to the kth class.

By the above, we can get the inference rules about classification of recognition image:

$$\begin{array}{ll} \text{known} & D_{i1}(u) \to P_{i1}(u) \\ & \cdots\cdots \\ & D_{im_i}(u) \to P_{im_i}(u) \\ \text{and given} & D_{ij}^*(x) \\ \hline \text{receive} & P_{ij}^*(x) \end{array} \quad , \tag{12}$$

where m_i is the number of samples corresponds to the ith class, $D_{ij}(u)$ is the distance for element u to the jth sample of ith class, $P_{ij}(u)$ is the probability.

If we divide the identified image into c categories, the details of the new algorithm are described as follows:

Step 1. Calculate the mean vector m_{ij} and covariance matrix C_{ij} of each sample by Eq.(5).

Step 2. Using the following equation to compute the Euclidean distance $D_{ij}^*(x)$ between the unknown elements x and m_{ij};

$$D_{ij}(x) = \sqrt{x(x - m_{ij})^T (x - m_{ij})} \qquad (i = 1, 2, \cdots, n) \ .$$

Step 3. Calculate $D_{ij}(u)$ and $P_{ij}(u)$ of the element u in sample block according to Eq.(5).

Step 4. Integrating the results by steps (2) and (3), derive from E_y which satisfies the previous conditions by Eq.(11). And then calculate the probability $P_{ij}^*(x)$ of the unknown element x belongs to the pattern class which corresponds to the sample block.

Step 5. In accordance with the principle of maximum membership, getting the probability $P_i^*(x)$ for x belonging to each pattern class;

$P_i^*(x) = \max\{P_{i1}^*(x), P_{i2}^*(x), \cdots, P_{im_i}^*(x)\}$.

Step 6. For the probability $P_i^*(x)$, assort the identified elements on the basis of the criteria of identification.

5 Experimental Results

In order to test the validity and applicability of the new algorithm, images which are got by low-altitude remote sensing system from Xiamen Passenger Station are used (image resolution 700×585). Next the comparison of the results' accuracy is finished among minimum distance criterion, traditional Bayes algorithm and the new algorithm respectively. And compares the results' accuracy of minimum distance criterion and traditional Bayes algorithm with the new algorithm respectively.

Experiment 1. Using minimum distance criterion, traditional Bayesian criterion and fuzzy entropy triple I algorithm respectively to get the water area in figure 3(a), and figure 3(b),(c) is the result of water area (where the white area shall be water) for minimum distance criterion and traditional Bayes algorithm respectively, figure 3(d) is the new algorithm's results. Through table 1, we have a conclusion that the new algorithm reduces the false consciousness number of assorting result better, which showing it has a good performance against false knowledge.

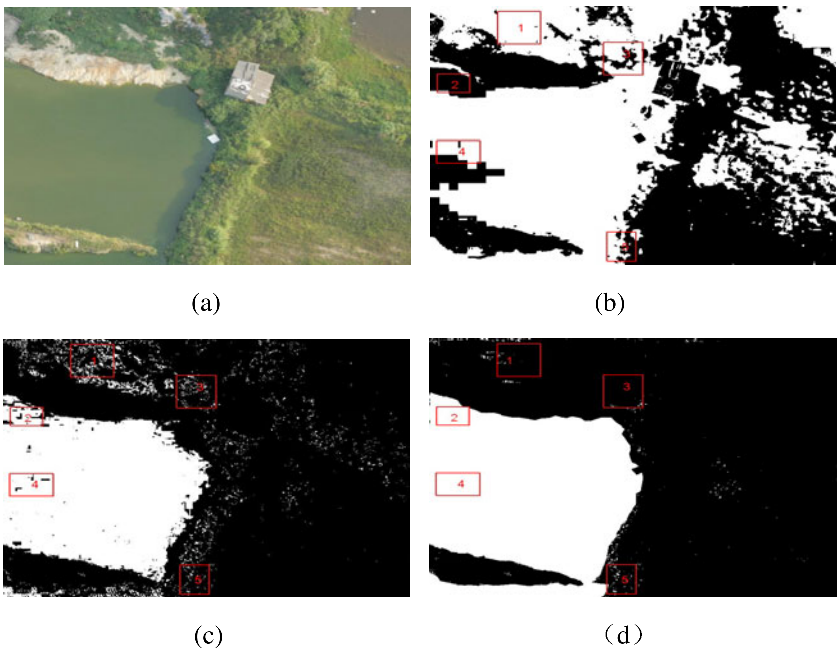


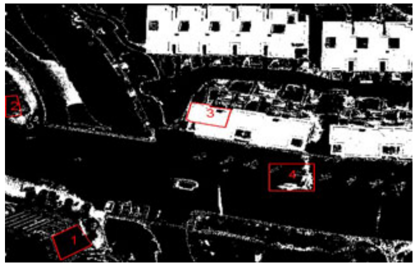
Fig. 3. The result of the proposed algorithm was compared to the other methods: (a) original RGB image; (b) the result of water area by using minimum distance discrimination; (c) the result of traditional Bayes algorithm; (d) the result of the water region by using the fuzzy entropy triple I algorithm;

Table 1. The accuracy of different methods

Methods	Area 1	Area 2	Area 3	Area 4	Area 5	Accuracy
The minimum distance criterion	0.0078	0.0604	0.2700	0.7599	0.4387	0.2954
Traditional Bayes algorithm	0.8000	0.8320	0.9231	0.9615	0.9529	0.8910
The new algorithm	0.9934	0.9976	0.9927	1.0000	0.9520	0.9881

Experiment 2. Using minimum distance criterion, traditional Bayes criterion and fuzzy entropy triple I algorithm respectively to get the buildings in figure 4(a), and figure 4(b),(c) and (d) are the results. The new algorithm can get very good results from table 2. Thus, assorting in different colors, the new algorithm has better applicability.

(a)



(b)

(c)

(d)

Fig. 4. The results on buildings of the three algorithms: (a) original RGB image; From (b) to (d): result of minimum distance discrimination method, traditional Bayes algorithm and fuzzy entropy triple I algorithm.

Table 2. The accuracy of different algorithms

Algorithms	Area 1	Area 2	Area 3	Area 4	Accuracy
The minimum distance criterion	0.0056	0.0482	0.9658	0.7990	0.5753
Traditional Bayes algorithm	0.0881	0.0138	0.9025	1.0000	0.6592
The new algorithm	0.5381	0.6284	0.9475	1.0000	0.8444

6 Conclusion

This paper proposes fuzzy entropy triple I algorithm, which has a combination of minimum distance theory and Bayesian probability theory, and shows good performance. Directly assorting the remote sensing images to high-resolution true color (RGB), the new algorithm can overcome the defect of collecting a single sample on the minimum distance discrimination method and traditional Bayes algorithm. The results of experiments show that the new algorithm for classification has higher precision, lower false accept rate and false reject rate, and stronger applicability for overcoming the limitations of traditional Bayes discrimination.

Of course, this algorithm can't recognize heterogeneous which has the same color, so that it needs to be further done with texture discrimination knowledge, this article is no longer considered because of limitations.

Acknowledgements. This work was supported by the National Natural Science Foundation of Fujian Province (No.A0810014), the key project on Department of Education of Fujian Province (No.JK2009017), and Innovation Fund projects of the younger on Science and Technology Department of Fujian Province (No.2009J05009).

References

1. Bian, Z., Zhang, X.: Pattern Recognition. Tsinghua University Press, Beijing (2001)
2. Chen, S., Li, J., Wang, X.: Fuzzy set theory and its application. Science Press (2005)
3. Li, F., Hongfan, S., Yong, Y., et al.: The Technology of Image Classification on Bayesian Classifier. Journal of Changchun University of Technology (Natural Science) 32(1), 132–134 (2009)
4. Wang, G.: Non-classical Mathematical Logic and Approximate Reasoning, 2nd edn. Science Press (2008)
5. Fang-fang, G., Tu-Yun, C., Zun-Quan, X.: Triple I methods for Fuzzy Reasoning Based on Maximum Fuzzy Entropy Principle. Fuzzy Systems and Mathematics 17(4), 55–59 (2003)
6. Laviolette, M., Seaman, J.W., Barrett, J.D., et al.: A probabilistic and statistical view of fuzzy methods. Technometrics 37(3), 249–261 (1995)
7. Fei, Y.: Probability representation and approximation abilities of fuzzy systems constructed by fuzzy entropy triple I method. Journal of Hainan Normal University (Natural Science) 21(4), 472–477 (2008)
8. Bhutani, K.R., Battou, A.: An application of fuzzy relations to image enhancement. Pattern Recognition Letters 16, 901–909 (1995)

Formalized Solutions for Reverse Triple-I under Theory of Truth Degree

Dong-li Liu

College of Chengyi, Jimei University, Fujian, Xiamen 361021, China
e-mail: liudl@jmu.edu.cn

Abstract. The problems of fuzzy reasoning reverse triple-I are proposed in n-valued propositional logic $Ł_n$ and ${L_n}^*$ .And truth degree formalized solutions for reverse triple-I problems are proved. Therefore, formal reasoning system of reverse triple-I method comes into being in n-valued systems $Ł_n$ and ${L_n}^*$.The work in the present paper lays a logical foundation for reverse triple-I methods of fuzzy reasoning.

Keywords: Truth degree, Reverse triple-I problems, Truth degree formalized solutions.

1 Introduction

It is well know that, fuzzy control has widely been applied to many industry and scientific research fields. Fuzzy inference is the key part of fuzzy control. Early in 1973,Compositional Rule of Inference (CRI) method was instituted by Zadeh1.Then, CRI method was widely used in fuzzy control. As Wang pointed out, CRI method lacks solid logic foundation, moreover, Wang proposed well-know Triple I algorithm by combining fuzzy logic and fuzzy inference to solve the problem of fuzzy modus Ponens(FMP) 2.Since then, a variety of research papers related to Wang's monograph have been published such as [5-9].One of these is Song and Wu's work9,they proposed Reverse triple I method from how to design fuzzy reasoning ruler. It's basic idea can be summarized as follows:For $A \in\ F(X)$,$B \in\ F(Y)$,and $A^* \in\ F(X)(B^* \in\ F(Y))$,its purpose is to seek a maximum $B^* \in\ F(Y)$(a minimum $A^* \in\ F(X)$) such that $(A^*(x) \to B^*(y)) \to (A(x) \to B(y))$ has the maximal possible value whenever $x \in X$ and $y \in Y$,where $F(X)$ and $F(Y)$ denote, respectively, the collections of all fuzzy subsets of X and Y.Diverse monographs related to Reverse triple I method had been accomplished10-12. Noticed that, the present paper is related to, but different from the above mentioned works. It aims to institute formal reasoning system of reverse triple

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 483–490.
springerlink.com

I problems. As Wang pointed out in[7], $G\ddot{o}$del system and product system are not suitable for establishing fuzzy logic. So, the formal version in the present paper is considered only in n-valued systems Ł$_n$ and $L_n{}^*$.

2 Preliminaries

2.1 n-Valued Logic Systems: Ł$_n$ and $L_n{}^*$

The set of formulas is a free algebra of the type $(\neg,\vee,\rightarrow)$,where S=$\{p_1$,p_2,$\ldots\}$ is a set of atomic formulas,$\neg$is an unary operation,$\vee$ and $\rightarrow$ are binary operations. n-valued propositional logic systems Ł$_n$ and $L_n{}^*$ are got by confining [0,1] to n-valued $L_n = \{0, \frac{1}{n-1}, \ldots, \frac{n-2}{n-1}\}$ respectively in Łukasiewicz and L^* systems. A homomorphism $v : F(S) \rightarrow L_n$ of type $(\neg,\vee,\rightarrow)$ is called a valuation of Ł$_n$ or $L_n{}^*$. The set of all valuations in Ł$_n$ and $L_n{}^*$ will be denoted by $\Omega(L_n)$ and Ω_n respectively[7]. We will denote them all by $\Omega(L_n)$ in the present paper for convenience sake. MV-algebra and R_0 -algebra are respective the corresponding algebraic system of formal systems (briefly Łuk) and L^*. We know $a \otimes b = (a \rightarrow b')'$ *and* $a \oplus b = a' \rightarrow b$ either in MV-algebra or in R_0-algebra M, where $\otimes$ and $\oplus$ are t-norm and t-remain norm [7], $'$ is defined by $a' = 1 - a, a, b \in M$.

We use $\rightarrow_L, \otimes_L$ and $\oplus_L$ to note Łuk implication operator, corresponding t-norm and t-remain norm in MV-algebra, and $\rightarrow_0, \otimes_0$ *and* $\oplus_0$ to note R_0 implication operator, corresponding t-norm and t-remain norm in R_0-algebra, then we have [7]:

(1) In MV-algebra:

$a \rightarrow_L b = (a' + b) \wedge 1, a \otimes_L b = (a + b - 1) \vee 0, a \oplus_L b = (a + b) \wedge 0.$

(2) In R_0 -algebra:

$$a \rightarrow_0 b = \begin{cases} 1, & a \leq b \\ a' \vee b, & a > b \end{cases}, a \otimes_0 b = \begin{cases} a \wedge b, & a + b > 1 \\ 0, & a + b \leq 1 \end{cases},$$

$$a \oplus_0 b = \begin{cases} 1, & a + b \geq 1 \\ a \vee b, & a + b < 1 \end{cases}$$

Let $A, B \in F(S)$. Define

$A \otimes B = \neg(A \rightarrow \neg B), A \oplus B = \neg A \rightarrow B.$

Then it is easily to check that: $A \otimes B = \neg(\neg A \oplus \neg B)$; $A \oplus B = \neg(\neg A \otimes \neg B)$.

Obviously, the next lemma is easy to verify:

Lemma 2.1.1. *Let $A, B \in F(S)$.*
(1) If $v \in \Omega_n$, then $v(A \otimes B) = v(A) \otimes_0 v(B), v(A \oplus B) = v(A) \oplus_0 v(B)$.
(2) If $v \in \Omega(L_n)$, then $v(A \otimes B) = v(A) \otimes_L v(B), v(A \oplus B) = v(A) \oplus_L v(B)$.

Moreover, if $v(A) = v(B) for\ all\ v \in \Omega_n$, then A and B are called logically equivalent, denoted as $A \approx B$[7].

2.2 About the Theory of Truth Degree

Let $A(p_1, p_2, \ldots, p_m)$ be a formula containing m atoms formula $p_1, p_2, \ldots, p_m$. Then A lead to a function $\bar{A}(x_1, x_2, \ldots, x_m)$, which is obtained by connecting the variables $x_1, x_2, \ldots, x_m$ in [0,1] with the operators $\neg$,$\vee$, and $\rightarrow$ on [0,1] in the same way as A is constructed from $p_1, p_2, \ldots, p_m$.

Definition 2.2.1[7]. *Let $A(p_1, p_2, \ldots, p_m)$ be a formula containing m atoms formula $p_1, p_2, \ldots, p_m$ in n-valued system $Ł_n$ (or $L_n{}^*$). Then the truth degree of A is defined by*

$\tau(A) = \frac{1}{n^m} \sum_{i=1}^{n-1} \frac{i}{n-1} \mid \bar{A}^{-1}(\frac{i}{n-1}) \mid,$

where $\mid \bar{A}^{-1}(\frac{i}{n-1}) \mid$ is the number of elements in the set $\bar{A}^{-1}(\frac{i}{n-1})$.

Lemma 2.2.2[7]. *Let $A, B \in F(S)$. Then*
(1) A is a tautology iff $\tau(A) = 1$; A is a contradiction iff $\tau(A) = 0$.
(2) If $A \approx B$, then $\tau(A) = \tau(B)$.

The next lemma is easily to prove:

Lemma 2.2.3. (1) *If $v(A) \le v(B)$ for $\forall v \in \Omega_(n)$, then $\tau(A) \le \tau(B)$.*
(2) *If there exists $v_0 \in \Omega_(n), v_0(A) < 1$, then $\tau(A) < 1$.*
(3) *If $v(A) \ge \alpha$ for $\forall v \in \Omega_(n)$, then $\tau(A) \ge \alpha$.*

We denote tautologies by $\bar{1}$, and contradictions by $\bar{0}$.

There are some symbols, ideals and results immediately used in the present paper, please refer to [7].

3 Formalized Solutions for Problems of Reverse Triple I

Definition 3.1. *Let $A, B \in F(S)$. Define*

$$A \le B \; iff \; v(A) \le v(B) \; for \; \forall v \in \Omega_(n),$$

then $(F(S), \le)$ is a set of pre-order on $F(S)$, and we call B is larger than A.

Problem 1 Reverse Triple IMP Let $A, B, A^* \in F(S)$, $A \rightarrow B$ *and* A^* be given. Then $B^* \in F(S)$ is the largest formula of $(F(S), \le)$, which lead $\tau((A^* \rightarrow B^*) \rightarrow (A \rightarrow B))$ to the maximum.

Problem 2 Reverse Triple IMT Let $A, B, B^* \in F(S)$, $A \rightarrow B$ *and* B^* be given. Then $A^* \in F(S)$ is the smallest formula of $(F(S), \le)$, which lead $\tau((A^* \rightarrow B^*) \rightarrow (A \rightarrow B))$ to the maximum.

The problems of Reverse Triple IMP and Reverse Triple IMT are all called Reverse Triple I problems.

Remark 1. *It is clearly that, $B^*(A^*)$ in problem 1 (problem 2) is the formulas, which lead the truth degree $\tau((A^* \rightarrow B^*) \rightarrow (A \rightarrow B))$ to the maximum. That is, the solutions of problem 1 (problem 2) are sought under the theory of truth degree. Therefore, we call $B^*(A^*)$ truth degree formalized solutions.*

Remark 2. *Obviously, if* $A \to B$ *is a tautology i.e.,* $v(A \to B) = 1$ *for* $\forall v \in \Omega_(n)$*, then* $\tau((A^* \to B^*) \to (A \to B)) \equiv 1$*. Furthemore truth degree formalized solutions of problem 1 is* $B^* \equiv \bar{1}$ *and truth degree formalized solutions of problem 2 is* $A^* \equiv \bar{0}$*. In an ordinarly way, we only consider formalized solutions in the case of* $\tau(A \to B) < 1, \tau(B^*) < 1$ *and* $\tau(B^*) > 0$*.*

If we don't emphasize specially in the following part, "$\to$" notes either $\to_0$ or $\to_L$, $\otimes$ and $\oplus$ is similarly.

Lemma 3.2. *In formal systems* $Ł_n$ *or* ${L_n}^*$*,* $\tau((A^* \to B^*) \to (A \to B)) \leq \tau(A^* \oplus (A \to B)$*, where* $A, A^*, B, B^* \in F(S)$*.*

Proof. Owing to $\forall v \in \Omega_(n)$, $v(A^* \to \bar{0}) \leq v(A^* \to B^*)$, so $v((A^* \to B^*) \to (A \to B)) = v(A^* \to B^*) \to (A \to B) \leq v(A^* \to \bar{0}) \to v(A \to B)$.

In formal system $Ł_n$, $v(A^* \to \bar{0}) \to v(A \to B) = (v(A^*) \to 0) \to v(A \to B) = v'(A^*) \wedge 1 \to v(A \to B) = v(\neg A^* \to (A \to B)) = v(A^* \oplus (A \to B))$.

In formal system ${L_n}^*$, $v(A^* \to \bar{0}) \to v(A \to B) = v'(A^*) \vee 0 \to v(A \to B) = v(A^* \oplus (A \to B))$.

Thus in $Ł_n$ and ${L_n}^*$, $v((A^* \to B^*) \to (A \to B)) \leq v(A^* \oplus (A \to B))$ for $\forall v \in \Omega_(n)$. Furthermore, $\tau((A^* \to B^*) \to (A \to B)) \leq \tau(A^* \oplus (A \to B)$.

Remark 3. *Lemma 3.2 tell us that in formal systems* $Ł_n$ *and* ${L_n}^*$*, if* $A^*, A \to B$ *are given, then the maximum value of* $\tau((A^* \to B^*) \to (A \to B))$ *is the truth degree of the formula* $A^* \oplus (A \to B)$*.*

Proposition 3.3. *Let* $A, B, A^* \in F(S)$*. If* $v(A^*) + v(A \to B) \geq 1$ *for* $\forall v \in \Omega_(n)$*, then the truth degree formalized solutions of Reverse triple IMP are* $B^* \approx A^* \otimes (A \to B)$*, which satisfies* $\tau((A^* \to B^*) \to (A \to B)) = 1$*.*

Proof. It follows from the condition of this proposition that $\forall v \in \Omega_(n)$, $v(A^* \oplus (A \to B)) = v(A^*) \oplus v(A \to B) = v'(A^*) \to v(A \to B)$. Notice that $v'(A^*) \leq v(A \to B)$, thus $v'(A^*) \to v(A \to B) = 1$,that is, $v(A^* \oplus (A \to B)) = 1$. Hence $\forall v \in \Omega_(n)$,$v(A^* \oplus (A \to B)) = 1$, therefore $\tau((A^* \oplus (A \to B)) = 1$, it follows from Lemma 3.2 that the maximum value of $\tau((A^* \to B^*) \to (A \to B))$ is 1.

(1) Firsterly, we prove that B^* satisfies $\tau((A^* \to B^*) \to (A \to B)) = 1$. In fact, it follows from $B^* \approx A^* \otimes (A \to B)$ that $\forall v \in \Omega_(n)$, $v(B^*) \leq v(A^*), v(B^*) \leq v(A \to B)$. Notice that $v'(A^*) \leq v(A \to B)$, then.

Case 1. If $\to$ is $\to_0$, then it follows from $v(A^*) \to v(B^*) = v'(A^*) \vee v(B^*) \leq v(A \to B)$ that $(v(A^*) \to v(B^*)) \to v(A \to B) = 1$, hence $v((A^* \to B^*) \to (A \to B)) = 1$,so $\tau((A^* \to B^*) \to (A \to B)) = 1$.

Case 2. If $\to$ is $\to_L$, then $v(B^*) \leq v(B^*) = v(A^* \otimes (A \to B)) = v(A^*) \otimes_L v(A \to B)$. It follows from $v(A^*) + v(A \to B) \geq 1$ that $v(A^*) \otimes_L v(A \to B) = (v(A^*) + v(A \to B) - 1) \vee 0 = v(A^*) + v(A \to B) - 1$, that is, $v(B^*) \leq v(A^*) + v(A \to B) - 1$, so $v(A \to B) \geq 1 - v(A^*) + v(B^*) = (v'(A^*) + v(B^*)) \wedge 1 = v(A^*) \to_L v(B^*)$, hence $\forall v \in \Omega_(n)$, $v((A^* \to B^*) \to (A \to B)) = (v(A^*) \to v(B^*)) \to v(A \to B) = 1$, that is $\tau((A^* \to B^*) \to (A \to B)) = 1$.

(2) Next, we will prove that B^* is the largest formula satisfying $\tau((A^* \to B^*) \to (A \to B)) = 1$. Suppose that $C^* \in F(S)$, and $B^* < C^*$, then $A^* \otimes (A \to B) < C^*$. Furthermore $\forall v \in \Omega_(n)$, $v(A^*) \otimes v(A \to B) = v(A^* \otimes (A \to B)) < v(C^*)$ therefore $v(A \to B) < v(A^*) \to v(C^*)$, hence $\tau((A^* \to B^*) \to (A \to B)) < 1$.

Remark 4. *Let's suppose that* $\exists v_0 \in \Omega_(n), v_0(A^*) + v_0(A \to B) < 1$.*Then in formal system* $Ł_n$*, we have* $v_0(A^* \oplus (A \to B)) = v_0(A^*) \oplus_L v_0(A \to B) = (v_0(A^*) + v_0(A \to B)) \wedge 0 \leq v_0(A^*) + v_0(A \to B) < 1$.

And in formal system L_n^*, $v_0(A^* \oplus (A \to B)) = v_0(A^*) \oplus_0 v_0(A \to B) = v_0(A^*) \vee v_0(A \to B) < 1$. Hence $\tau((A^* \to B^*) \to (A \to B)) \leq \tau(A^* \oplus (A \to B)) < 1$. That is, there is not a $B^* \in F(S)$, which satisfies $\tau((A^* \to B^*) \to (A \to B)) = 1$. Hence the conclusion of proposition 3.3 is not found. That is, the condition $\forall v \in \Omega_(n)$, $v(A^*) + v(A \to B) \geq 1$ in proposition 3.3 is sufficient and necessary. We have

Proposition 3.4. *In* L_n^* *system, let* $A, B, A^* \in F(S)$*. If* $v(A^*) + v(A \to B) < 1$ *and* $v(A^*) > \frac{1}{2}$ *for* $\forall v \in \Omega_n$*, then the truth degree formalized solutions of Reverse triple IMP are* $B^* \approx A^* \wedge \neg A^*$*, which satisfies* $\tau((A^* \to B^*) \to (A \to B)) = \tau(A^*)$.

Proof. It follows from the condition in this proposition that $\forall v \in \Omega_n$, $v(A^* \oplus (A \to B) = v(\neg A^* \to (A \to B)) = v'(A^*) \to_0 v(A \to B) = v(A^*) \vee v(A \to B)$. Notice that $v(A^*) > \frac{1}{2}$,then $v(A \to B) < \frac{1}{2}$, so $v(A^* \oplus (A \to B)) = v(A^*) \vee v(A \to B) = v(A^*)$. Hence $A^* \oplus (A \to B) \approx A^*$. It follows from Lemma 2.2.2 that $\tau(A^* \oplus (A \to B)) = \tau(A^*)$. Then it follows from Lemma 3.2 that the maximum value of $\tau((A^* \to B^*) \to (A \to B))$ *is* $\tau(A^*)$.

(1) We first prove that B^* satisfies $\tau((A^* \to B^*) \to (A \to B)) = \tau(A^*)$. We have $\forall v \in \Omega_n$, $v((A^* \to B^*) \to (A \to B)) = v(A^* \to (A^* \wedge \neg A^*)) \to v(A \to B) = v'(A^*) \vee (v(A^*) \wedge v'(A^*)) \to v(A \to B) = v'(A^*) \to v(A \to B) = v(A^*) \vee v(A \to B) = v(A^*)$. Then $(A^* \to (A^* \wedge \neg A^*)) \to (A \to B) \approx A^*$. It is easy to know that $\tau((A^* \to (A^* \wedge \neg A^*)) \to (A \to B)) = \tau(A^*)$.

(2) Next, we will prove that B^* is the largest formula. Suppose $C^* \in F(S), B^* < C^*$, that is $A^* \wedge \neg A^* < C^*$, then $\forall v \in \Omega_n, v(A^*) \wedge v'(A^*) < v(C^*)$. Hence $v((A^* \to C^*) \to (A \to B)) = (v(A^*) \to v(C^*)) \to v(A \to B) < (v(A^*) \to v(A^*) \wedge v'(A^*)) \to v(A \to B) = v'(A^*) \to v(A \to B) = v(A^*) \vee v(A \to B) = v(A^*)$. It follows from Lemma 2.2.3 that $\tau((A^* \to C^*) \to (A \to B)) < \tau(A^*)$, so $B^* \approx A^* \wedge \neg A^*$ is the largest formula that satisfies $\tau((A^* \to B^*) \to (A \to B)) = \tau(A^*)$.

Remark 5. *From* $\forall v \in \Omega_n, v(A^*) > \frac{1}{2}$ *we know* $\tau(A^*) > \frac{1}{2}$*. So it is rational for this proposition supposing this condition because of belief beyond half principle.*

As you know,we only consider system $\overset{*}{L_n}$ in Proposition 3.4, but how to $Ł_n$ system?

Proposition 3.5. *In L_n system, let $A, A^*, B \in F(S)$. If $v(A^*) + v(A \to B) < 1$ for $\forall v \in \Omega(L_n)$, then the truth degree formalized solutions of Reverse triple IMP are $B^* \equiv \bar{0}$.*

Proof. We have from Lemma 3.2 that the maximum value of $\tau((A^* \to B^*) \to (A \to B))$ *is* $\tau(A^* \oplus (A \to B))$. Notice that, in the condition of $\forall v \in \Omega(L_n)$ one hand, $v((A^* \to B^*) \to (A \to B)) = (v(A^*) \to v(B^*)) \to v(A \to B) = (v'(A^*) + v(B^*)) \wedge 1 \to v(A \to B) = ((v'(A^*) + v(B^*)) \to v(A \to B)) \vee v(A \to B) = (v(A^*) + v(A \to B) - v(B^*)) \vee v(A \to B)$. On the other hand,$v(A^* \oplus (A \to B)) = v(\neg A^* \to (A \to B)) = v'(A^*) \to v(A \to B) = (v(A^*) + v(A \to B)) \wedge 1 = v(A^*) + v(A \to B)$.

Suppose that $\tau((A^* \to B^*) \to (A \to B))$ can get the maximum value, then at least one of the following two conditions holds:

(1) $v(A^*) + v(A \to B) - v(B^*) = v(A^*) + v(A \to B)$.

$v(A \to B) = v(A^*) + v(A \to B)$.

(2) We have from (1) that $v(B^*) = 0$, then $B^* \equiv \bar{0}$. And from (2) we know $v(A^*) = 0$, which is not consistent with our preceding supposition. Then Proposition 3.2.4 holds.

We will discuss Reverse triple IMT in the following part.

Proposition 3.6. *Let $A, B^*, B \in F(S)$. If $v(B^*) \leq v(A \to B)$ for $\forall v \in \Omega_{(n)}$, then the truth degree formalized solutions of Reverse triple IMT are $A^* \approx B^* \oplus \neg(A \to B)$, which satisfies*

$$\tau((A^* \to B^*) \to (A \to B)) = 1.$$

Proof. (1) We first prove that $A^* \approx B^* \oplus \neg(A \to B)$ satisfies $\tau((A^* \to B^*) \to (A \to B)) = 1$.

Case 1. In formal system Ł$_n$.

Notice that $v(A^*) = v(B^*) \oplus v'(A \to B) = v'(B^*) \to v'(A \to B) = v(A \to B) \to v(B^*)$,we have $\forall v \in \Omega(L_n), v((A^* \to B^*) \to (A \to B)) = (v(A^*) \to v(B^*)) \to v(A \to B) = ((v(A \to B) \to v(B^*)) \to v(B^*)) \to v(A \to B) = v(A \to B) \vee v(B^*) \to v(A \to B)$. It follows from $v(B^*) \leq v(A \to B)$ that we know the value of the above equality is equal to 1.

Case 2. In formal system L_n^*.

Notice that $\forall v \in \Omega_n, v((A^* \to B^*) \to (A \to B)) = (v(A^*) \to v(B^*)) \to v(A \to B) = v'(A^*) \vee v(B^*) \to v(A \to B)$, then from $v'(A^*) \leq v(A \to B), v(B^*) \leq v(A \to B)$ we know $v((A^* \to B^*) \to (A \to B)) = 1$. Hence $A^* \approx B^* \oplus \neg(A \to B)$ satisfies $\tau((A^* \to B^*) \to (A \to B)) = 1$.

(2) Next we will prove that A^* is the least formula. It is easy to verify that $(A^* \to B^*) \to (A \to B) \approx (\neg B^* \to \neg A^*) \to (A \to B)$. (*)

Notice that $v(B^*) \leq v(A \to B)$ for $\forall v \in \Omega(n)$ then $v(\neg B^*) + v(A \to B) \geq 1$. From Proposition 3.3, we know that $\neg A^* \approx \neg B^* \otimes (A \to B)$ is the largest formula satisfying $\tau((A^* \to B^*) \to (A \to B)) = 1$. Therefore, $A^* \approx B^* \oplus \neg(A \to B)$ is the least formula satisfying $\tau((A^* \to B^*) \to (A \to B)) = 1$.

Proposition 3.7. *In formal system L_n^*, let $A, B^*, B \in F(S)$. If $v(B^*) > v(A \to B)$ and $v(B^*) < \frac{1}{2}$ for $\forall v \in \Omega_n$, then the truth degree formalized solutions of Reverse triple IMT are $A^* \approx B^* \vee \neg B^*$, which satisfies*

$$\tau((A^* \to B^*) \to (A \to B)) = 1 - \tau(B^*).$$

Proof. From $(A^* \to B^*) \to (A \to B) \approx (\neg B^* \to \neg A^*) \to (A \to B)$, we know $\tau((A^* \to B^*) \to (A \to B)) = \tau((\neg B^* \to \neg A^*) \to (A \to B))$. Notice the conditions, we have $\forall v \in \Omega_n, v(\neg B^*) + v(A \to B) < 1$ and $v(\neg B^*) > \frac{1}{2}$, then it follows from Proposition 3.4 that the maximum value of $\tau((A^* \to B^*) \to (A \to B))$ *is* $\tau(\neg B^*) = 1 - \tau(B^*)$. It is easy to know that $A^* \approx B^* \vee \neg B^*$ satisfies $\tau((A^* \to B^*) \to (A \to B)) = 1 - \tau(B^*)$. In addition, we know $\tau((A^* \to B^*) \to (A \to B)) = \tau((\neg B^* \to \neg A^*) \to (A \to B)) = 1 - \tau(B^*)$, where the largest formula is $\neg A^* \approx \neg B^* \wedge \neg(\neg B^*) = \neg B^* \wedge B^*$, from Proposition 3.4 that the least formula is $A^* \approx B^* \vee \neg B^*$.

Remark 6. *Similar to proposition 3.3, the condition $v(B^*) \leq v(A \to B)$ for $\forall v \in \Omega_(n)$ in Proposition 3.6 is sufficient and necessary.*

Remark 7. *The remainders notice that in proposition 3.2.6,there is a condition of $v(B^*) < \frac{1}{2}$ for $\forall v \in \Omega_n$, here B^* is later part of $A^* \to B^*$.From the result $A^* \approx B^* \vee \neg B^*$,we know that $v(A^*) > \frac{1}{2}$ for $\forall v \in \Omega_n$ so the condition is reasonable.*

It is easy to verify the following proposition.

Proposition 3.8. *Let $A, B^*, B \in F(S)$. In formal system $Ł_n$, if $v(B^*) > v(A \to B)$ for $\forall v \in \Omega(Ł_n)$, then the truth degree formalized solutions of Reverse triple IMT are $A^* \equiv \bar{1}$.*

4 Conclusion

In Reverse triple I method provides a new and valid way for how to design more reasonable fuzzy reasoning regular. In the present paper, the problems of fuzzy reasoning reverse triple I are defined. Then, based on the thought of graded, we discuss reverse triple IMP and reverse triple IMT. The truth degree formalized solutions of reverse triple I are given. The graded method presented in the present paper lead the algorithgmic realization of solution. It is easy to verify that, our conclusions can be, in a sense, brought into line with Song's reverse triple I method. Therefore, the present paper lays a logical foundation for reverse triple I methods of fuzzy reasoning.

Acknowledgements. This work described here is supported by the project from the scientific research fund of Sichuan provincial education department (No. 08Za132).

References

1. Zadeh, L.A.: Outline of a new approach to the analysis of complex systems and decision processes. IEEE Trans. Systems Man Cybernet. 3, 28–44 (1973)
2. Wang, G.J.: The triple I method with total inference rules of fuzzy reasoning. Science in China E 29, 43–53 (1999)
3. Wang, G.J., Fu, L.: Unified forms of triple I method. Comput. Math. Appl. 49, 923–932 (2005)
4. Pei, D.W.: Fullimplication triple I algorithms and their consistency in fuzzy reasoning. Journal Mathematical Research Exposition 24, 359–368 (2004)
5. Song, S., Feng, C., Lee, E.S.: Triple I method of fuzzy reasoning. Comput. Math. Appl. 44, 1567–1579 (2002)
6. Wang, G.J.: Non-classia Mathematical Logic and Approximate Reasoning. Science Publisher, Beijing (2000)
7. Wang, G.J.: Introduction to Mathematical Logic and Resolution Principle (The second Publishing). Science in China Press, Beijing (2006)
8. Wang, G.J.: Triple I method and interval value fuzzy reasoning. Science in China (series E) 30, 331–340 (2000)
9. Song, S.J., Wu, C.: Reverse Triple I method of Fuzzy Reasoning. Science in China (Series F) 45, 331–340 (2002)
10. Song, S.J., Wu, C.: Reverse restricted triple I method of Fuzzy Reasoning. Progress in Natural Science 12, 95–100 (2002)
11. Qin, K.Y., Pei, Z.: Opposite Direction Triple Method under Lukasiewicz Implication Operator. Fuzzy Systems and Mathematics 19, 1–5 (2005)
12. Peng, J.Y.: Opposite Direction Triple Method under common Implication Operators. Progress in Natural Science 15, 404–410 (2005)
13. Liu, D.L., Wang, G.J.: Formal Solution to the Problem of -triple I in Many-valued Logic System. Computer Engineering and Applications 41, 51–53 (2005)

Properties of L-Extremally Disconnected Spaces

Ji-shu Cheng[1] and Shui-li Chen[2]

[1] School of Information Engineering, Hangzhou Dianzi University, Hangzhou 310018, China
e-mail: chengjishu@126.com

[2] School of Science, Jimei University, Xiamen361021, P.R. China
e-mail: sgzx@jmu.edu.cn

Abstract. The concept of L-extremally disconnected spaces is introduced and investigated in this paper, which is the generalization of the concept of fuzzy extremally disconnected spaces due to Ghosh. In L-extremally disconnected spaces, it is proved that the concepts of semi-open, pre-open and alpha-open sets are uniform. We will also show that two theorems are incorrect in Ghosh's [1] article by the counterexample.

Keywords: L-topological space, extremally disconnected space, pre-open set, semi-pre-open set.

1 Introduction and Preliminaries

In 1992, Ghosh generalized the notion of extremally disconnected spaces to fuzzy topological spaces(see, [1]). In [8], Park and Lee introduced and studied the notion of fuzzy extremally disconnected spaces in fuzzy bitopological spaces. In this paper, we define the concept of L-extremally disconnected spaces and discuss its properties using the semi-open (semi-closed), pre-open(pre-closed) and α-open (α-closed). Using counterexamples shows that some results in [1] are false.

Throughout this paper, L denotes a completely distributive lattice with order-reversing involution $'$. X denotes a nonempty set, L^X denotes the sets of all L-sets on X, while (L^X, δ) denotes an L-fuzzy topological space (briefly, L-space). The elements in δ are called L-open sets(briefly, open sets) and the elements in $\delta' = \{B \mid B' \in \delta\}$ are called L-closed sets(briefly, closed sets). An element a of L is said to be $\vee$-irreducible or a molecule if $a \leq b \vee c$ implies that $a \leq b$ or $a \leq c$, where $b, c \in L$. The set consisting of all nonzero $\vee$-irreducible elements of L will be denoted by M, and the set consisting of all nonzero $\vee$-irreducible elements of L^X will be denoted by $M(L^X)$,

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 491–498.
springerlink.com

i.e., $M(L^X) = \{x_\alpha | x \in X, \alpha \in M\}$. A^-, A° and A' will denote the closure, the interior and the pseudo-complement of $A \in L^X$, respectively. Let $\eta(x_\alpha) = \{P \in \delta' | x_\alpha \not\leq P\}$ and the elements in $\eta(x_\alpha)$ are said to be R-neighborhoods of x_α[9]. $A \in L^X$ is called regular open[2] (semi-closed [3]) sets iff $A = A^{-\circ}$ (there exists an open set $B \in \delta$ such that $B \leq A \leq B^-$). If A is a regular open (semi-open) set, then A' is called regular closed (semi-open) set. All regular open, regular closed, semi-open and semi-closed sets are denoted by $\mathrm{Ro}(L^X)$, $\mathrm{Rc}(L^X)$, $\mathrm{So}(L^X)$ and $\mathrm{Sc}(L^X)$, respectively. Let $A \in L^X$. Then $A^-_\theta = \bigwedge\{C \in \delta' | A \leq C^\circ\}$, $A^-_\delta = \bigwedge\{C \in \mathrm{Ro}(L^X) | A \leq C\}$ and $A^-_s = \bigwedge\{C \in \mathrm{Sc}(L^X) | A \leq C\}$ are called the θ- closure[4, 6], δ-closure[2, 5] and S-closure[3], respectively. It is obvious that $A \leq A^-_s \leq A^- \leq A^-_\delta \leq A^-_\theta$.

Theorem 1.1. Let (L^X, δ)be L-space and $x_\alpha \in$M(L^X). Then
(1)[9] $x_\alpha \in A^-$ iff for each $P \in \eta(x_\alpha)$, $A \not\leq P$.
(2)[3] $x_\alpha \in A^-_s$ iff for each $P \in S\eta(x_\alpha)$, $A \not\leq P$.
(3)[4] $x_\alpha \leq A^-_\theta$ iff for each $P \in \eta(x_\alpha)$, $A \not\leq P^\circ$.
(4)[5] $x_\alpha \in A^-_\delta$ iff for each $P \in \delta\eta(x_\alpha)$, $A \not\leq P$.

2 Some Properties of α-open, Pre-open and Semi-pre-open Sets

In this section, we discuss the concepts of generalized open (closed) sets in L-spaces. Some nice results of these generalized open (closed) are obtained.

Definition 2.1. *Let (L^X, δ)be an L-space and $A \in L^X$.*
(1) A is called pre-open (semi-pre-open) iff $A \leq A^{-\circ}$ ($A \leq A^{-\circ-}$). The family of all pre-open (semi-per-open)sets will be denoted by $Po(L^X)$ ($Spo(L^X)$).
(2) A is called pre-closed (semi-pre-closed) iff $A^{\circ-} \leq A$ ($A^{\circ-\circ} \leq A$). The family of all pre-closed (semi-pre-closed) sets will be denoted by $Pc(L^X)$ ($Spc(L^X)$).
(3) A is called α-open if $A \leq A^{\circ-\circ}$ for each $A \in L^X$.
(4) A is called α-closed if $A^{-\circ-} \leq A$ for each $A \in L^X$.
$\alpha o(L^X)$ and $\alpha c(L^X)$ will denote the families of all α-open (α-closed) sets in (L^X, δ), respectively.
(5) $A^-_p = \bigwedge\{D \in Pc(L^X) | A \leq D\}$ ($A^-_{sp} = \bigwedge\{D \in Spc(L^X) | A \leq D\}$).
(6) $B^\circ_p = \bigvee\{G \in Po(L^X) | G \leq B\}$ ($B^\circ_{sp} = \bigvee\{G \in Spo(L^X) | G \leq B\}$).
(7) $A^-_\alpha = \bigwedge\{D \in \alpha c(L^X) | A \leq D\}$, $A^\circ_\alpha = \bigvee\{G \in \alpha o(L^X) | G \leq A\}$.

Let (L^X, δ)be an L-space and $x_\alpha \in M(L^X)$. We write $s\eta(x_\alpha) = \{P \in \mathrm{Sc}(L^X) | A \not\leq P\}$, $p\eta(x_\alpha) = \{Q \in \mathrm{Pc}(L^X) | A \not\leq Q\}$, $sp\eta(x_\alpha) = \{Q \in \mathrm{Spc}(L^X) | A \not\leq Q\}$ and $\alpha\eta(x_\alpha) = \{P \in \alpha c(L^X) | A \not\leq P\}$, respectively.

Remark 2.1. *Every open set is α-open and every α-open set is a semi-open as well as pre-open set.*

Theorem 2.1. *Let (L^X, δ)be an L-space. Then*
(1) $(A_s^\circ)^- = (A^-)_s^- = A^-$ and $(A^\circ)_s^\circ = (A_s^\circ)^\circ = A^\circ$ for each $A \in L^X$.
(2) $A^{-\circ} \leq A_s^-$ and $(A_s^-)^\circ = A^{-\circ}$ for each $A \in L^X$.
(3) $A_s^\circ \leq A^{\circ -}$ and $(A_s^\circ)^- = A^{\circ -}$ for each $A \in L^X$.

Proof. One can easily check that (1) and (2). We only prove (3). Since A_s^- is a semi-closed set, there exists $C \in \delta'$ such that $C^\circ \leq A \leq C$, which implies $C^\circ \leq A_s^- \leq A^- \leq C$. Thus we have $C^\circ \leq A^{-\circ} \leq C^\circ$ and $C^\circ \leq A_s^-$. So $C^\circ = A^{-\circ} \leq A_s^-$ is proved. It is clear that $A^{-\circ} \leq (A_s^-)^\circ$ because $A^{-\circ} \leq A_s^-$ which was just proved. From $A_s^- \leq A^-$, we get $(A_s^-)^\circ \leq A^{-\circ}$ and hence $(A_s^-)^\circ = A^{-\circ}$.

Theorem 2.2. *Let (L^X, δ)be an L-space. Then*
(1) $A \in Sc(L^X)$ iff $A^\circ = A^{-\circ}$ for any $A \in L^X$.
(2) $A \in So(L^X)$ iff $A^- = A^{\circ -}$ for any $A \in L^X$.

Proof. (1) We first prove that $A \in \mathrm{Sc}(L^X)$ iff $A^{-\circ} \leq A$. In fact, suppose that A is in $\mathrm{Sc}(L^X)$. Then $A^{-\circ} \leq A_s^- = A$ because $A = A_s^-$ is semi-closed by Theorem 2.1(2). Conversely, assume that $A^{-\circ} \leq A$, then $A^{-\circ} \leq A \leq A^-$. we write $C = A^-$, then C is a closed set in $(L^X, \delta)$and satisfying $C^\circ \leq A \leq C$. Consequently, $A \in \mathrm{Sc}(L^X)$. From $A^{-\circ} \leq A$, we get $A^{-\circ} \leq A^\circ$. On the other hand, it is obvious that $A^\circ \leq A^{-\circ}$ and hence $A^\circ = A^{-\circ}$. Conversely, let $A^\circ = A^{-\circ}$. Since $A^\circ \leq A$, thus we have $A^{-\circ} \leq A$. Consequently, A is a semi-closed set, i.e., $A \in \mathrm{Sc}(L^X)$.

(2) Follows from (1).

Theorem 2.3. *Let (L^X, δ)be an L-space. Then*
(1) $A \in Po(L^X)$ iff $A_s^- = A^{-\circ}$.
(2) $A \in Po(L^X)$ iff A_s^- is a regular open set.
(3) $Ro(L^X)$=$Po(L^X) \bigcap Sc(L^X)$.

Proof. (1) If $A \in Po(L^X)$, then $A \leq A^{-\circ}$ by Definition 2.1, which implies $A_s^- \leq (A^{-\circ})_s^-$. Since $A^{-\circ}$ is semi-closed, thus $A_s^- \leq A^{-\circ}$ and hence $A_s^- = A^{-\circ}$ by Theorem 2.1(2). Conversely, it is clear.

(2) It follows from (1).

(3) Let $A \in \mathrm{Po}(L^X) \bigcap \mathrm{Sc}(L^X)$. Then $A = A_s^- = A^{-\circ}$ and so $A \in \mathrm{Ro}(L^X)$ by (1). Conversely, if $A \in \mathrm{Ro}(L^X)$, then $A = A^{-\circ}$. We have $A \leq A^{-\circ}$ and $A \geq A^{-\circ}$ which implies $A^\circ \geq A^{-\circ}$. So, $A \in \mathrm{Po}(L^X) \bigcap \mathrm{Sc}(L^X)$ by Theorem 2.2.

Theorem 2.4. *Let (L^X, δ)be an L-space. Then the following conditions are equivalent:*

(1) A is an α-open set.
(2) There exists an open set B such that $B \leq A \leq B^{-\circ}$.
(3) There exists a closed set Q such that $Q^{\circ -} \leq A' \leq Q$.
(4) A' is an α-closed set.

Proof. It is clear that (1) $\Longleftrightarrow$ (4) and (2) $\Longleftrightarrow$ (3). Now, we need prove only (1) $\Longleftrightarrow$ (2). Let A be α-open. Then $A \leq A^{\circ-\circ}$. We write $B = A^{\circ}$, then $B \leq A \leq B^{-\circ}$. Conversely, if there exists $B \in \delta$ such that $B \leq A \leq B^{-\circ}$. We have $B^{-} \leq A^{\circ -}$ by $B \leq A^{\circ}$ and hence $A \leq B^{-\circ} \leq A^{\circ-\circ}$. Consequently, A is α-open.

Theorem 2.5. *Let (L^X, δ)be an L-space and $A \in L^X$. Then*
(1) A is α-open iff A is both semi-open and pre-open.
(2) A is α-closed iff A is both semi-closed and per-closed.

Proof. (1) If A is α-open, then $A \leq A^{\circ-\circ} \leq A^{\circ-}$ as well as $A \leq A^{-\circ}$. So A is semi-open and pre-open. Conversely, if A is pre-open and semi-open, i.e., $A \leq A^{-\circ}$ as well as $A \leq A^{\circ-}$, then $A^{-} \leq A^{\circ-}$ and that $A^{-\circ} \leq A^{-}$, which implies $A \leq A^{-\circ} \leq A^{\circ-\circ}$. This shows that $A \leq A^{\circ-\circ}$ and hence A is α-open.

(2) A is α-closed, iff A' is α-open, iff A' is pre-open and semi-open, iff A is pre-closed and semi-closed.

Corollary 2.1. *Let (L^X, δ)be an L-space. Then*
(1) $A \in \alpha o(L^X)$ iff $A^{-} = A^{\circ-}$ and $A \leq A^{-\circ}$.
(2) $A \in \alpha c(L^X)$ iff $A^{\circ} = A^{-\ \circ}$ and $A^{\circ-} \leq A$.
(3) $\alpha o(L^X)$=$Po(L^X) \bigcap So(L^X)$.
(4) $\alpha c(L^X)$=$Pc(L^X) \bigcap Sc(L^X)$.

3 *L*-Extremally Disconnected Spaces

In this section, we generalize the notion of fuzzy extremally disconnected spaces to L-spaces and investigate its properties. We will show that $A^{-} = A^{-}_{s}$ and $A^{-} = A^{-}_{\theta}$ etc., for equivalent conditions of extremally disconnected spaces are incorrect.

Definition 3.1.*An L-space $(L^X, \delta)$is called L-extremally disconnected, if the closure of every open set is open.*

The following proposition is obvious.

Proposition 3.1. *In an L-space (L^X, δ), the following conditions are equivalent:*
(1) (L^X, δ) is L-extremally disconnected.
(2) Every regular closed set is open.
(3) $\forall A \in \delta$, $A^{-} \leq A^{-\circ}$.
(4) $\forall B \in \delta'$, $B^{\circ-} \leq B^{\circ}$.

Theorem 3.1. *Let (L^X, δ) be an L-extremally disconnected space. Then $A_\theta^- = (A_\theta^-)_\theta^-$ for every $A \in L^X$.*

Proof. For arbitrary $x_\alpha \in M(L^X)$ and $x_\alpha \not\leq A_\theta^-$, then there exists $P \in \eta(x_\alpha)$ such that $A \leq P^\circ$. Since P is closed and (L^X, δ) is extremally disconnected, $P^{\circ -} \leq P^\circ$. Thus we have $A^- \leq P^{\circ -} \leq P^\circ$ and $A_\theta^- \leq (A^-)_\theta^- \leq (P^\circ)_\theta^- = P^{\circ -} \leq P^\circ$. This shows that $x_\alpha \not\leq (A_\theta^-)_\theta^-$ and hence $A_\theta^- = (A_\theta^-)_\theta^-$.

Theorem 3.2. *An (L^X, δ) is L-extremally disconnected iff for each $A, B \in \delta$ with $A \leq B'$, then $A^- \leq B'^\circ$.*

Proof. Let A and B be open sets with $A \leq B'$. Since B' is closed, by $A \leq B'$, we have $A^- \leq B'$. Notice that A^- is open because (L^X, δ) is An L-extremally disconnected space, thus $A^- = A^{-\circ} \leq B'^\circ$. Conversely, assume that the condition of theorem holds. For each $A \in \delta$, we write $B = A'^\circ$, then $B \in \delta$ and $A \leq A^- = B'$. This proves that $A^- \leq B'^\circ = A^{-\circ}$. So (L^X, δ) is An L-extremally disconnected space by Proposition 3.1.

Lemma 3.1. *If (L^X, δ) is L-extremally disconnected, then $A^- = A_s^-$ for all $A \in So(L^X)$.*

Proof. For every $A \in \mathrm{So}(L^X)$, we need only prove that $A^- \leq A_s^-$. Let $x_\alpha \not\leq A_s^-$ and $x_\alpha \in M(L^X)$. Then there exists $Q \in S\eta(x_\alpha)$ such that $A \leq Q$. Since Q is a semi-closed set, there exists closed set P such that $P^\circ \leq Q \leq P$. By hypothesis, (L^X, δ) is L-extremally disconnected, $P^{\circ -} \leq P^\circ$ by Proposition 3.1, and hence $P^{\circ -} \in \eta(x_\alpha)$. Since A is semi-open, $A^- = A^{\circ -}$. By $A \leq Q$ and $P^\circ \leq Q \leq P$, we get $A \leq A^- \leq A^{\circ -} \leq Q^{\circ -} \leq P^{\circ -}$. So, $x_\alpha \not\leq A^-$. This proves that $A^- \leq A_s^-$ and hence $A^- = A_s^-$.

Remark 3.1. *$A^- = A_s^-$ for all $A \in So(L^X)$ only is necessary condition of fuzzy extremally disconnected space, but it is not sufficient condition. So, Theorem 5 in [1] is erroneous. We have the following counterexample.*

Example 3.1. *Let $L = X = [0, 1]$. Fuzzy sets A and B are defined as follows:*

$$A(x) = 0.7\ ,\ B(x) = 0.3 \quad \textit{for all } x \in X,$$

$$\delta = \{0,\ 1,\ B\}.$$

Then (L^X, δ) is fuzzy topological space. One easy seen $B^- = B' = A$. It is clear that $B \leq A \leq B^-$ and $A^- = A_s^-$, so A is semi-open. But (L^X, δ) is not fuzzy extremally disconnected space because $A^- = A'$ is not an open set in (L^X, δ).

Since every semi-open set is semi-pre-open set, by Example 3.1, Theorem 12 is also false in [1].

Lemma 3.2. *In an L-space (L^X, δ), $A^- = A^-_\delta$ for all $A \in So(L^X)$.*

Proof. If $x_\alpha \not\leq A^-_\delta$ and $x_\alpha \in M(L^X)$, then there exists $P \in \delta\eta(x_\alpha)$ such that $A \leq P$. Since P is regular closed and A semi-open set, $P = P^{\circ -}$ and $A^- = A^{\circ -}$. Thus we have $A \leq A^- = A^{\circ -} \leq P^{\circ -} = P$, which shows that $x_\alpha \not\leq A^-$ and hence $A^-_\delta \leq A^-$. On the other hand, $A^- \leq A^-_\delta$ is obvious and so $A^- = A^-_\delta$.

From Theorem 2.2, Lemma 3.1 and Lemma 3.2, the following results are obvious.

Theorem 3.3. *Let (L^X, δ) be an L-st. Then the following conditions are equivalent:*

(1) (L^X, δ) is An L-extremally disconnected space.
(2) A^- is open for every $A \in So(L^X)$.
(3) A^-_δ is open for every $A \in So\ (L^X)$.
(4) A^-_s is open for every $A \in So(L^X)$.
(5) A^- is open for every $A \in So(L^X)$.
(6) A° is closed for every $A \in Sc(L^X)$.

Theorem 3.4. *Let $(L^X, \delta)$is an L-space. For each $A \in Po(L^X)$. Then*

$$A^- = A^-_\delta = A^-_\theta.$$

Proof. It is obvious that $A^- \leq A^-_\delta \leq A^-_\theta$ for each $A \in \mathrm{Po}(L^X)$. Thus it remains to show that $A^-_\theta \leq A^-$. Now, suppose that $x_\alpha \not\leq A^-$ $(x_\alpha \in M(L^X))$. Then there exists $P \in \eta(x_\alpha)$such that $A \leq P$, which means that $A^- \leq P$ and $A^{-\circ} \leq P^\circ$. Since A is pre-open, by Definition 2.1, $A \leq A^{-\circ} \leq P^\circ$ and so, $x_\alpha \not\leq A^-_\theta$. Thus we have proved $A^-_\theta \leq A^-$.

Lemma 3.3. *Let (L^X, δ) be an L-space. Then $A^- = A^-_\delta$ for every $A \in Spo(L^X)$.*

Proof. To prove this, it is sufficient to prove that $A^-_\delta \leq A^-$ for every $A \in \mathrm{Spo}(L^X)$. Now, let $x_\alpha \in M(L^X)$ and $x_\alpha \not\leq A^-$. Then there exists $P \in \eta(x_\alpha)$ such that $A \leq P$, which implies $A^{-\circ -} \leq P^{\circ -}$. Since $P^{\circ -} \leq P$, $P^{\circ -}$ is a regular closed set and $A \in \mathrm{Spo}(L^X)$, thus we have $P^{\circ -} \in \delta\eta(x_\alpha)$ and $A \leq A^- = A^{-\circ -} \leq P^{\circ -}$. So, $x_\alpha \not\leq A^-_\delta$. This proves that $A^-_\delta \leq A^-$ and hence $A^- = A^-_\delta$.

By Theorem 3.3, Lemma 3.9 and Lemma 3.3, it is obvious that the following theorem.

Theorem 3.5. *In an (L^X,δ), then the following conditions are equivalent:*

(1) (L^X,δ) is L-extremally disconnected.
(2) A^- is open for every $A \in Spo(L^X)$.
(3) A^-_δ is open for every $A \in Spo(L^X)$.
(4) A^-_θ is open for every $A \in Po(L^X)$.
(5) A^- is open for every $A \in Po(L^X)$.

Theorem 3.6. *An (L^X,δ) is L-extremally disconnected iff $So(L^X) \subset \alpha o(L^X)$.*

Proof. Let $A \in \mathrm{So}(L^X)$. Then exists $Q \in \delta$ such that $Q \leq A \leq Q^-$. Since (L^X,δ) is L-extremally disconnected, $Q^- \in \delta$. Thus we get $Q \leq A \leq Q^{-\circ}$ and hence $A \in \alpha o(L^X)$ by Theorem 2.4. Conversely, let A be arbitrary regular closed set. Then $A \in \mathrm{So}(L^X)$ and by hypothesis, $A \in \alpha o(L^X)$. Thus we have $A \leq A^{\circ-\circ} \leq A^{-\circ}$ and $A = A^{-\circ}$ because A is closed. Consequently, (L^X,δ) is L-extremally disconnected by Proposition 3.1.

Theorem 3.7. *An L-space (L^X,δ) is L-extremally disconnected iff $So(L^X) \subset Po(L^X)$.*

Proof. It is similar with the proof of Theorem 3.6.

Since $\alpha o(L^X) \subset \mathrm{Po}(L^X)$ and $\alpha o(L^X) \subset \mathrm{So}(L^X)$, we have:

Corollary 3.1. *In an L-space (L^X,δ), the following statements are equivalent:*

(1) (L^X,δ) is An L-extremally disconnected space.
(2) $\alpha o(L^X)= \mathrm{Po}(L^X)$.
(3) $\alpha o(L^X)= \mathrm{So}(L^X)$.
(4) $\mathrm{So}(L^X)= \mathrm{Po}(L^X)$.

Definition 3.2. [3] *Let $f : (L^X,\delta)\longrightarrow (L_1^Y,\delta_1)$be an order homomorphism.*

(1) f is called semi-continuous if $f^{-1}(B) \in So(L^X)$ for each $B \in \delta_1$.
(2) f is called irresolute if $f^{-1}(B) \in So(L^X)$ for each $B \in So(L^X)$.
(3) f is called almost open if $f(A) \in \delta_1$ for each $A \in So(L^X)$.

Theorem 3.8. *If $f : (L^X,\delta) \longrightarrow (L_1^Y,\delta_1)$ is almost open as well as semi-continuous, then $f(B) \in Po(L_1^Y)$ for every $B \in Po(L^X)$.*

Proof. Suppose that B is in $\mathrm{Po}(L^X)$, then $f(B) \leq f(B_s^-) \leq (f(B))^-$ by Theorem1.1 in [3]. Thus we have $B_s^- \in \mathrm{Ro}(L^X)$ and $f(B_s^-) \in \mathrm{Po}(L^X)$ because f is almost open. Since $(f(B_s^-))_s^- = (f(B_s^-))^{-\circ}$ by Theorem 2.3, $(f(B))_s^- \leq (f(B_s^-))_s^- = (f(B_s^-))^{-\circ} \leq (f(B))^-$. By $f(B) \leq f(B_s^-) \leq (f(B))^-$, we get $(f(B))^{-\circ} \leq (f(B_s^-))^{-\circ} \leq (f(B))^{-\circ}$ and hence $(f(B))^{-\circ} = (f(B_s^-))^{-\circ}$. This shows that $f(B) \leq (f(B))_s^- \leq (f(B_s^-))^{-\circ} = (f(B))^{-\circ}$ and so $f(B) \in \mathrm{Po}(L^X)$.

Lemma 3.4. *If $f : (L^X, \delta) \longrightarrow (L_1^Y, \delta_1)$ is almost open as well as semi-continuous, then f is irresolute.*

Proof. Assume that $A \in \mathrm{Sc}(L_1^Y)$, then $A^{-\circ} \leq A$. Since f is semi-continuous, $f^{-1}(A^-) \in \mathrm{Sc}(L^X)$. Thus it is obvious by Theorem 2.2, $(f^{-1}(A^-))^{-\circ} = (f^{-1}(A^-))^{\circ}$ because $f^{-1}(A^-)$ is semi-closed. Since f is also almost open,

$$f((f^{-1}(A^-))^{\circ}) \leq (ff^{-1}(A^-))^{\circ} \leq A^{-\circ} \leq A,$$

which implies $(f^{-1}(A^-))^{\circ} \leq f^{-1}(A)$. Moreover, $(f^{-1}(A))^{-\circ} \leq (f^{-1}(A^-))^{-\circ} = (f^{-1}(A^-))^{\circ}$, so we get $(f^{-1}(A))^{-\circ} \leq f^{-1}(A)$ and hence $f^{-1}(A) \in \mathrm{Sc}(L^X)$. This shows that f is irresolute.

Theorem 3.9. *Let $f : (L^X, \delta) \longrightarrow (L_1^Y, \delta_1)$ be almost open surjection as well as semi-continuous. If (L^X, δ) is L-extremally disconnected, then so is (L_1^Y, δ_1).*

Proof. Let A be arbitrary semi-open in (L_1^Y, δ_1). Since f is semi-continuous and almost open, $f^{-1}(A)$ is irresolute by Lemma 3.4, so that $f^{-1}(A)$ is semi-open. By Theorem 3.6, we have $f^{-1}(A) \in \mathrm{Po}(L^X)$. Since f is semi-continuous and almost open surjection, $A = ff^{-1}(A) \in \mathrm{Po}(L^X)$ by Theorem 3.8. This proves $\mathrm{So}(L^X) \subset \mathrm{Po}(L^X)$, and so (L_1^Y, δ_1) is an L-extremally disconnected space by Theorem 3.6.

References

1. Ghosh, B.: Fuzzy extremally disconnected spaces. J. Fuzzy Sets and Systems 86, 245–250 (1992)
2. Cheng, J.S.: Some properties of δ-continuous order homomorphisms. Fuzzy Systems and Mathematics 4, 38–41 (1997)
3. Chen, S.L., Cheng, J.S.: The characterizations of semi-continuous and irresolute order homomorphisms on fuzzes. Fuzzy Sets and Systems 64, 105–112 (1994)
4. Chen, S.L., Cheng, J.S.: θ-Convergence of nets of L-fuzzy sets and its applications. J. Fuzzy Sets and Systems 86, 235–240 (1997)
5. Cheng, J.S.: On δ-continuous, δ-open and δ-closed order homomorphisms. J. Fuzzy Mathematics 4, 885–894 (1999)
6. Cheng, J.S., et al.: Regular Separation Axiom in L-topological Spaces. Chinese Journal of Engineering Mathematics 25, 181-184(2008)
7. Liu, Y.M., Luo, M.K.: Fuzzy topology. World Scientific Publishing, Singapore (1997)
8. Park, J.H., Lee, B.Y.: Fuzzy pairwise extremally spaces. Fuzzy Sets and Systems 98, 201–206 (1998)
9. Wang, G.J.: Theory of L-fuzzy topological space. Shaanxi Normal University Press, Xi'an (1988)
10. Wang, G.J.: Order homomorphisms on fuzzes. J. Fuzzy Sets and Systems 12, 281–288 (1984)

Fuzzy Cluster Analysis of Regional City Multi-level Logistics Distribution Center Location Plan

Yong-chang Ren[1], Tao Xing[2], Qiang Quan[3], and Guo-qiang Zhao[4]

[1] College of Information Science and Engineering,
Bohai University, Jinzhou, 121013, China
e-mail: rycryc@sina.com
[2] Beijing Research Center of Urban Systems Engineering,
Key Laboratory of City Operation and Management,
Beijing, 100089, China
e-mail: xingtao@mail.tsinghua.edu.cn
[3] Electronic & Information Engineering College,
Liaoning Technology University, Jinzhou, 121001, China
e-mail: Quanqiang1982@126.com
[4] College of Electrical and Control Engineering,
Liaoning Technical University, Fuxin 123000, China
e-mail: zhaoguoqiang@163.com

Abstract. Establish the cities multi-level logistics distribution center in the region to improve the flow efficiency and economic benefits, enhance regional competitiveness, promote the rational allocation of regional resources and effective use has an important role. This article use fuzzy clustering analysis method to study. First discusses the preparations of fuzzy clustering analysis. includes the construction of influence Factors indicator system, determine the factor weight set, quantify the value of level 2 indicators, calculate the value of level 1 indicators, etc. Then study the fuzzy clustering analysis method and steps, including data standardization, the establishment of fuzzy relations, fuzzy clustering, etc. Finally, through a specific location planning instance to illustrate the whole process of cluster analysis. The results show that, the method of fuzzy cluster analysis Provide a scientific method to the regional city multi-level logistics distribution center location plan, but in the practical application also needs integrated more factors to make a final decision.

Keywords: City logistics centre, region multi-level logistics distribution, location plan, fuzzy cluster analysis.

1 Introduction

Regional economy is the economic union which combined by the factors, such as natural, economic, ethnic, cultural traditions and social development, is the Specialization of social economic activities and cooperation reflected in the space. Regional logistics and regional economy is the interdependent unity, is the main

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 499–508.
springerlink.com

elements of the regional economy [1]. Regional logistics as an important part of the regional economic activity, is the strong pillars of playing the regional function, also is the leading force of the formation and development of regional economic system. In the region, establish the cities multi-level logistics distribution center is playing an active promoting role ,such as in improving production and circulation efficiency and economic benefits, reflecting the reasonable allocation of regional resources and effective utilization, improving competitiveness in regional markets and changing the distribution of manufacturing enterprises and production mode.

Clustering is the process that according to the similarity of things to distinguish between and classify the process, cluster analysis is using mathematical method to study and deal with the classification of the given object [2]. Application of ordinary mathematical methods to the classification of clustering method known as ordinary cluster analysis, also called hard classifying, each object to be identified should be strictly divided into a class, this classification marked clearly. Actually, most objects have intermediary in nature and generic respect, application of fuzzy mathematical methods to analyze the cluster analysis known as the fuzzy cluster analysis, also called soft division, fuzzy set theory provide powerful analysis tools to this soft classifying, people began to use fuzzy approach to process the clustering problem. In the Research field of logistics, fuzzy cluster analysis has been applied to product classification, logistics performance measurement, logistics facilities location, etc. City is the core carrier of the regional logistics, the commodity distribution and processing center, is the application object of the clustering. Through the comparison of the conditions on the horizontal inter-city, establish the hierarchy, determine the level of logistics center, form the regional Multi-level distribution center network.

2 Fuzzy Clustering Analysis Preparation

Preparing influence factors and data value for the cluster analysis. The Preparation process includes the construction of influence factors indicator system and data value of each index. When the indicator system is in multi-level, need to establish the factors weight set of each level, through the weight set calculating the lowest level indicator data up to the higher level, until 1uantization to the 1 Level indicator. The present study build 2 level index system, just need calculating the 2 level index data up to the 1 level index data by the weight set.

A. Construction of Influence Factors Indicator System

The city logistics distribution center location is an integrated decision-making problem which involve many influence factors, in the location process all factors have the influence in different degrees, only consider the integration of all factors can make the urban logistics distribution center location decisions in a more rational and more scientific. Generally, when make the city distribution center location decision should major consider the center construction economic factors, social factors, infrastructure, natural environment and business environment, etc. The references [3-5] has already studied the city distribution center location influence factors index system in depth. Based on summarizing these studies, combined with the experience of writers and experts, summarize the influence Factors indicator system as showed in Table 1.

Table 1. City logistics distribution centre location influence factors indicator system

target layer	criteria layer (level 1 indicators)	index layer (level 2 indicators)	
Optimal location for city logistics distri-bution centre Plan U	economic factors U_1	land prices	u_{11}
		construction costs	u_{12}
		transport costs	u_{13}
		operational costs	u_{14}
	social factors U_2	industrial policy	u_{21}
		human resources	u_{22}
		environmental protection	u_{23}
		resident attitudes	u_{24}
	infrastructure U_3	road traffic	u_{31}
		circumstances condition	u_{32}
		public facilities	u_{33}
		waste disposal	u_{34}
	natural environment U_4	weather conditions	u_{41}
		geological conditions	u_{42}
		hydrological conditions	u_{43}
		topography conditions	u_{44}
	business environment U_5	competitors	u_{51}
		market demand	u_{52}
		service level	u_{53}
		product features	u_{54}

B. Determine the Weight Factors Set

In the influence factors indicator system, the degrees of each factor weight is different. To show the importance of each factor, each factor u_i should be given appropriate weight w_i. the set formed by the weights is called the factor weights set, short called the weights set.

The level 1 indicators weights set expressed as:

$$W = (W_1, W_2, \cdots, W_n) . \tag{1}$$

In the last formula, n is the number of the level 1 indicators.

The level 2 indicators weights set expressed as:

$$W_i = (w_{i1}, w_{i2}, \cdots, w_{im}) . \tag{2}$$

In the last formula $(i=1,2,\cdots n)$ express the level 1 indicators, m express the quantity of level 2 indicators corresponding to the appropriate level 1 indicators.

C. Quantify the Level 2 Indicators

The set formed by the various evaluation results of evaluation objects, expressed as [6]:

$$V=\{v_1,v_2,\cdots,v_n\}\ . \tag{3}$$

Based on integrated considering all influence factors, get the optimal evaluation results from the evaluation set.

To each influence factors of level 2 indicators, determine the membership degree of the evaluation object relative to the evaluation set elements. Need to establish fuzzy mapping [6] from U to $F(V)$:

$$\begin{aligned} &f:U\rightarrow F(V),\forall u_i\in U,u_i\mapsto \underset{\sim}{f}(u_i) \\ &=\frac{r_{i1}}{v_1}+\frac{r_{i2}}{v_2}+\cdots+\frac{r_{im}}{v_m}=\sum_{j=1}^{m}\frac{r_{ij}}{v_j} \end{aligned}\ . \tag{4}$$

In this formula r_{ij} express the membership that u_i belongs to v_j .

Membership, is also called membership function or fuzzy relation coefficient, is the key to describe the object fuzziness. Membership function is the objective measure of the subjective factors for the fuzzy object, on being given time trying to minimize the influence of subjective factors, furthest reflect the fuzzy objective characteristics. Commonly used methods are fuzzy statistical method, dualistic contrast compositor method, distribution method, etc.

Express the results of quantifying the level 2 indicators:

$$R_i=(r_{i1},r_{i2},\cdots r_{im})\ . \tag{5}$$

D. Calculate the Level 1 Indicators

According to level 2 indicators value and weight [8], calculate the level 1 indicators. The value of the level 1 indicator is obtained by the weighted calculation many level 2 indicators of it's, expressed as following formula:

$$c_i=W_i^T\bullet R_i^T=(w_{i1},w_{i2},\cdots,w_{im})\bullet(r_{i1},r_{i2},\cdots r_{im})^T\ . \tag{6}$$

On the previous formula c_i express the i-th level 1 indicator, the value of i and j value ibid.

Level 1 indicators value can express by matrix:

$$C=\begin{bmatrix} C_1 \\ C_2 \\ \mathrm{M} \\ C_i \\ \mathrm{M} \\ C_m \end{bmatrix}=\begin{bmatrix} c_{11} & c_{12} & \Lambda & c_{1j} & \Lambda & c_{1n} \\ c_{21} & c_{22} & \Lambda & c_{2j} & \Lambda & c_{2n} \\ \mathrm{M} & \mathrm{M} & \Lambda & \mathrm{M} & \Lambda & \mathrm{M} \\ c_{i1} & c_{i2} & \Lambda & c_{ij} & \Lambda & c_{in} \\ \mathrm{M} & \mathrm{M} & \Lambda & \mathrm{M} & \Lambda & \mathrm{M} \\ c_{m1} & c_{m2} & \Lambda & c_{mj} & \Lambda & c_{mn} \end{bmatrix}\ . \tag{7}$$

On the previous formula $(i=1,2,\cdots m)$, m express the distribution centre location number; $(j=1,2,\cdots n)$, n express the level 1 indicators number; c_{ij} express the j-th level 1 value indicator value of the i-th distribution centre.

3 Fuzzy Clustering Analysis Procedure

Fuzzy cluster analysis procedure can be summarized as: data standardization, fuzzy relations establishment, fuzzy clustering [6][7].

A. Data Standardization

In practical problems, the different data may have different dimension, to make different dimensions of data can be compared, and the data need proper transformation.

Set the clustering object as $A_1,A_2,\cdots,A_n$, $U=\{A_1,A_2,\cdots,A_n\}$ is the simple set, considerations factors (or called sample indicator) are $B_1,B_2,\cdots,B_m$, A_i can be described by m data, set the data corresponding to Ai is $(x'_{i1},x'_{i2},\cdots,x'_{im})$ $(i=1,2,\cdots m)$, can measure the n data $(x_{1k},x_{2k},\cdots,x_{nk})$ $(k=1,2,\cdots m)$ to Bk. The data standardization formulas as follows:

$$x''_{ik}=\frac{x'_{ik}-\bar{x}'_k}{S_k}. \tag{8}$$

On the previous formula, $\bar{x}'_k$ is the average value of the k-th indicator is:

$$\bar{x}'_k=\frac{1}{n}\sum_{i=1}^{n}x'_{ik}, \tag{9}$$

S_k is the standard deviation of the k-th indicator is:

$$S_k=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x'_{ik}-\bar{x}'_k)^2} \qquad \begin{array}{l} i=1,2,\cdots,n \\ k=1,2,\cdots,m \end{array}. \tag{10}$$

After the conversion, the average value of each variable is 0, the standard deviation is 1, and can eliminate the influence of dimension, but not necessarily in interval [0,1]. According to the requirements of fuzzy matrix, the data should be compressed in closed interval [0, 1], use extreme transformation formula:

$$x_{ik}=\frac{x''_{ik}-x''_{\min k}}{x''_{\max k}-x''_{\min k}}. \tag{11}$$

On the previous formula, $x''_{\min k}$ is the minimum of x''_k; $x''_{\max k}$ is the maximum of x''_k.

After the conversion of extreme value, there is $0\le x_{ik}\le 1$, get the standard array $(x_{i1},x_{i2},\cdots,x_{im})$ $(i=1,2,\cdots n)$.

B. Fuzzy Relations Establishment

Fuzzy relations establishment is also called calibration, is calculated the similarity coefficient r_{ij} $(i, j = 1,2,\cdots n)$ between the object classification, to be similar to the matrix $R = (r_{ij})_{n\times n}$, is fuzzy relations.

Set universe of discourse $U = \{A_1, A_2, \cdots, A_n\}$, each factor of the universe of discourse is a sample, each sample is m-dimensional vector, $(x_{i1}, x_{i2}, \cdots, x_{im})$, there are many methods to calculate the r_{ij} . There are three methods most commonly used as follows:

(1) Arithmetic average of the minimum method

$$r_{ij} = \frac{\sum_{k=1}^{m}(x_{ik} \wedge x_{jk})}{\frac{1}{2}\sum_{k=1}^{m}(x_{ik} + x_{jk})} . \tag{12}$$

(2) Geometric average of minimization method

$$r_{ij} = \frac{\sum_{k=1}^{m}(x_{ik} \wedge x_{jk})}{\sum_{k=1}^{m}\sqrt{(x_{ik} + x_{jk})}} . \tag{13}$$

(2) Correlation coefficient method

$$r_{ij} = \frac{\sum_{k=1}^{m}\left|x_{ik} - \bar{x}_i\right|\left|x_{jk} - \bar{x}_j\right|}{\sqrt{\sum_{k=1}^{m}(x_{ik} - \bar{x}_i)^2}\sqrt{\sum_{k=1}^{m}(x_{jk} - \bar{x}_j)^2}} . \tag{14}$$

On the previous formula, $\bar{x}_i = \frac{1}{m}\sum_{k=1}^{m} x_{ik}$, $\bar{x}_j = \frac{1}{m}\sum_{k=1}^{m} x_{jk}$.

C. Fuzzy Clustering

According to the calibration established the fuzzy matrix R . Generally speaking, its nature is reflexivity and symmetry, and does not satisfy transitivity, just a fuzzy similarity matrix, only if R is a fuzzy equivalence matrix to cluster, so R need to be transformed into equivalence fuzzy matrix R .

About the equivalence fuzzy matrix, there is Theorem 1 in [7]: Set R is a reflexive and symmetric relations of $U = \{A_1, A_2, \cdots, A_n\}$, that R is $n*n$ fuzzy similarity matrix, then there exists a minimum natural number $k(k \le n)$, make R^k as a fuzzy similarity matrix, and constant presence $R^w = R^k$ to the natural number which is greater than k. R^k is called the transitive closure matrix of R , marked as $t(R)$.

According to the previous theorem, the $n*n$ fuzzy similarity matrix R can be transformed into equivalence fuzzy matrix $t(R)$ by the transitive closure. From the fuzzy matrix R , demand in turn squared: $R \to R^2 \to R^4 \to \cdots$, when appear

$R^k \bullet R^k = R^k$ the first time, shows R^k has the nature of transfer, R^k is the transitive closure which is the demand of $t(R)$.

After the R is transformed into equivalence matrix R^k, do interception on the appropriate limit value, to obtain the required classification.

4 Planning Instance

A region is formed by 5 cities, such as X_1, X_2, X_3, X_4, X_5, establishing a logistics centre in each city. According to the different influence factors, to establish the different level logistics centre, preliminary decision to establish the level 1 and level 2 logistics centre, level 1 provide the service to level 2. determine the logistics centre level classification with the fuzzy cluster analysis method.

A. Quantify the Value of Level 2 Indicators

According to the method of quantifying the value of level 2 indicators, the result is expressed as Table 2:

B. Calculate the Value of Level 1 Indicator

According to the part of the index weight that is given in the references [3-5], combined with the experience of writers and experts, use AHP to determine the value of weight as follows:

$$A_1 = (a_{11}, a_{12}, a_{13}, a_{14}) = (0.32, 0.20, 0.38, 0.10),$$
$$A_2 = (a_{21}, a_{22}, a_{23}, a_{24}) = (0.33, 0.41, 0.15, 0.11),$$
$$A_3 = (a_{31}, a_{32}, a_{33}, a_{34}) = (0.32, 0.16, 0.42, 0.10),$$
$$A_4 = (a_{41}, a_{42}, a_{43}, a_{44}) = (0.25, 0.22, 0.18, 0.35),$$
$$A_5 = (a_{51}, a_{52}, a_{53}, a_{54}) = (0.30, 0.41, 0.07, 0.22).$$

According to the formula (6), the value of level 1 indicator is as follows:

C. Data Standardization and the Establishment of Fuzzy Relations

According to the formulas (8), (9), (10), (11), after making the data standardization, according to Formula (12), the establishment of fuzzy relations as follows:

$$R = \begin{bmatrix} 1 & 0.85 & 0.92 & 0.63 & 0.77 \\ 0.85 & 1 & 0.85 & 0.63 & 0.77 \\ 0.92 & 0.85 & 1 & 0.63 & 0.77 \\ 0.63 & 0.63 & 0.63 & 1 & 0.63 \\ 0.77 & 0.77 & 0.77 & 0.63 & 1 \end{bmatrix}.$$

Table 2. Quantify the value of level 2 indicators

level 2 indicators	symbol	X_1	X_2	X_3	X_4	X_5
land prices	u_{11}	76	63	68	72	88
construction costs	u_{12}	60	56	82	49	64
transport costs	u_{13}	69	76	84	65	73
operational costs	u_{14}	57	64	73	56	63
industrial policy	u_{21}	85	87	79	92	90
human resources	u_{22}	65	74	71	86	39
environmental protection	u_{23}	38	45	61	58	47
resident attitudes	u_{24}	52	56	85	90	73
road traffic	u_{31}	56	67	38	91	76
circumstances condition	u_{32}	80	62	71	88	35
public facilities	u_{33}	71	58	38	82	49
waste disposal	u_{34}	36	81	62	75	45
weather conditions	u_{41}	28	72	58	87	71
geological conditions	u_{42}	65	71	92	80	69
hydrological conditions	u_{43}	66	61	57	56	58
topography conditions	u_{44}	73	54	71	96	54
competitors	u_{51}	57	49	41	51	56
market demand	u_{52}	78	79	65	86	75
service level	u_{53}	46	48	49	71	35
product features	u_{54}	37	91	82	67	48

Table 3. The value of Level 1 indicators

level 1 indicators	symbol	X_1	X_2	X_3	X_4	X_5
economic factors	U_1	68	66	77	63	75
social factors	U_2	68	71	73	84	37
Infrastructure	U_3	64	63	45	85	55
natural environment	U_4	58	63	69	83	62
business environment	U_5	60	70	60	70	60

D. Fuzzy Cluster

To the fuzzy relation R in 4.4, by $r_{ii} = 1 \quad (i = 1,2,\cdots,n)$ know that R is reflexive; by $r_{ij} = r_{ji} \quad (i, j = 1,2,\cdots,n)$ know that R is symmetry. According to theorem 1, R is fuzzy equivalence matrix.

1) When $\lambda \geq 1$

Only diagonal elements to be equal or greater than 1, so all the diagonal elements to be transformed into 1, other elements to be transformed into 0, become a unit matrix, divided into 5 categories: $\{X_1\},\{X_2\},\{X_3\},\{X_4\},\{X_5\}$, each element is a class, and it is the most detailed classification.

2) When $\lambda \geq 0.92$

The element which is less than 0.92 to be transformed into 0, and the one which is greater than 0.92 to be transformed into 1, divided into 4 categories: $\{X_1, X_3\}$, $\{X_2\},\{X_4\},\{X_5\}$.

3) When $\lambda \geq 0.85$

There are 3 categories: $\{X_1, X_2, X_3\}$, $\{X_4\}$, $\{X_5\}$.

4) When $\lambda \geq 0.77$

There are 2 categories: $\{X_1, X_2, X_3, X_5\}$, $\{X_4\}$.

5) When $\lambda \geq 0.63$

There is only 1 category: $\{X_1, X_2, X_3, X_4, X_5\}$, it is crudest category.

The fuzzy clustering map based on the previous fuzzy clustering, is expressed as figure 1.

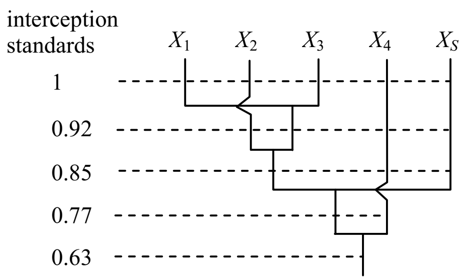


Fig. 1. Fuzzy clustering map

According to the result of category, when $\lambda = 0.77$, the alternative distribution centre address is divided into two categories, accordingly, we can get the initial location program of the region multi-level logistics distribution centre , that is planning the $\{X_4\}$ as level 1 distribution centre, planning $\{X_1, X_2, X_3, X_5\}$ as level 2 distribution centre, constitute a 2-levels distribution network.

5 Conclusion

The logistics centre location planning is the most important issue of the overall construction and development of the regional logistics network, the fuzzy clustering method is a scientific and practical method to solve the problem. The classification of the city logistics centre can promote the rational layout of the distribution logistics network, this rationality should not only adapt the current regional economic development, but also to adapt to future development needs. the fuzzy clustering emphasis on quantitative description, however, in practice, some influence factors can not be included in the indicators system to quantify, therefore, the results obtained by this method only as an important basis for decision making, the final result, still need be integrated more factors to make the final decision.

Acknowledgments. This work is supported by Chinese Natural Foundations (70871067), Liaoning Natural Foundations (20072207), Liaoning Higher Education School Creative Team's Founds (2008T090) and Liaoning doctoral Foundations (20091034).

References

1. Xiaolan, Z.: Fuzzy Clustering for Logistics Park Network Layout. Shanghai JiaoTong University, Shanghai (2007)
2. Leizhi, J.: The TPL Provider Selection Based on Fuzzy Clustering Analysis. Logistics Management 27(11), 34–37 (2004)
3. Jie, Y.: Logistics Center Location Research Based on Yibin. Zhejiang Gongshang University Hangzhou, Hangzhou (2009)
4. Yaosheng, J.: The Study of Location of Logistics Centers. Jiangsu University, Zhenjiang (2005)
5. Ling, W., Huifang, L.: Fuzzy Comprehensive Evaluation of the Logistics Center Location. Railway Freight Transport 28(11), 5–6 (2008)
6. Qizong, W.: System Engineering. Beijing Institute of Technology Press, Beijing (2005)
7. Hongji, L.: Fuzzy mathematics basis and pragmatic algorithm. Science Press, Beijing (2005)
8. Fenggen, L., Yaorong, C.: Fuzzy Cluster Analysis in Multilevel Distribution Center Location Application. Railway Transport and Economy 29(3), 59–61 (2007)
9. Taniguchi, E., Thompson, R.G., Yamada, T.: Modeling city logistics. Institute of Systems Science Research 35(2), 77–82 (1999)
10. Shen, Z.-J.M., Coull, C.: A joint location-inventory model. Transportation Science 37(1), 40–55 (2003)
11. Huifeng, G., Ligong, Q.: Location Decision Model of Multi-level Logistics Distribution Center and Computer Solution. Journal of Yibin University 30(6), 20–22 (2008)

The Establishment and Reliability Analysis of the Majorized Model of a Kind of Fuzzy Multidimensional Synthetical Decision

Xiao-nan Xiao

Xiamen University Tan Kah Kee College,
Zhangzhou 363105
e-mail: xiaoxn@xujc.com

Abstract. Applying the majorized method of fuzzy transformation and fuzzy integral, this paper discusses theoretically and pragmatically about how to make Fuzzy optimal multidimensional synthetical decision and analyzes its reliability, provides the optimal mathematical model of Fuzzy optimal multidimensional synthetical decision, promotes the expertise of synthetical decision and widens its application in fields of natural science and social science. This paper provides optimal decision for scientific management of teaching quality, reasonable employment flow of college graduates, modernized medical diagnosis and other complicated cases, which proves that the method this paper proposes is an ideal decision making approach that ensures a satisfactory result once it has been put into practice.

Keywords: Optimal multidimensional synthetical decision, majorized model, fuzzy transformation, fuzzy integral, possibility measure, strong law of large numbers, reliability analysis.

1 Introduction

With the development of modern science, its emphasis now shifts more and more rapidly from the research on the definite object by the method of analysis to the research on the indefinite object by the method of synthesis. After every concrete science has the typical phenomenon of either-or in its own sphere fully studied, it is now engaged in enlarging scope sphere and is ready to make the research on the untypical phenomenon of both-and. The trend of the penetrating between the different natural science, between the different social science, and between the social science and natural science appears apparently with the time going by. The former branch-bound line is broken and the frontier science springs up. The fuzzy mathematics appearing in 1960's is a great breakthrough in the prolongation of mathematics research.

We can use fuzzy mathematics to study the fuzzy phenomenon in the objective world, so the limitation of the tradition mathematics has been smoothed away and many problems which can not be solved by tradition mathematics have been solved. And

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 509–517.
springerlink.com

fuzzy mathematics applies a good situation to the mathematicization of natural science, social science, system science, idea science and body science. In recent years, many achievements concerning the fuzzy mathematics have appeared, but fuzzy decision theory is now in its childhood. So this article tries to conduct the research on the best fuzzy multidimensional synthetical decision and enlarge the sphere in which the best fuzzy multidimensional synthetical decision can be applied [1].

2 Research on BFMSD

Many systems in the world are influenced by many factors as well as have many aims, this is so-called "Multidimensional Synthetical Decision". Generally, a decision problem is always coupled with fuzziness, randomness and experience characteristics.

So it is common to adopt the fuzzy synthetical decision. The following would lay the emphasis on this problem.

In order to evaluate something synthetically, we must pay attention to the following three factors:

(1) Collection of elements $U=\{u_1,u_2,\cdots,u_n\}$ u_i shows the factor that has to be considered by something;

(2) Decision collection $V=\{v_1,v_2,\cdots,v_m\}$ v_i shows the stage of decision;

(3) Single elements decision, it is a fuzzy mapping from $U \to V$. According to the fuzzy mapping law, a fuzzy mapping $\underset{\sim}{f}$ can decide a fuzzy relation $\underset{\sim}{R}\,\underset{\sim}{f}$, it can be expressed by a fuzzy matrix $\underset{\sim}{R} \in M_{n\times m}$. So $\underset{\sim}{R}$ can be considered as a fuzzy transformation from $U \to V$.

Thus an evaluation space $(U,V,\underset{\sim}{R})$ may make up a model of synthetical evaluation.

Suppose a fuzzy subset in *U*,

$\underset{\sim}{A}=(a_1,a_2,\cdots,a_n)$, in which a_i represents the weighted number and it satisfies

$$\sum_{i=1}^{n} a_i = 1.$$

In the given fuzzy transformation R and factor weight A, we can get a fuzzy subset from the fuzzy relations compositive operation, that ~~is~~:

$$\underset{\sim}{A} \circ \underset{\sim}{R} = \underset{\sim}{B} \in M_{l\times m},$$

where $\mu_{\underset{\sim}{B}} = \mu_{\underset{\sim}{A}\circ\underset{\sim}{R}} = \bigvee_{u\in U}\{\mu_{\underset{\sim}{A}}(u) \wedge \mu_{\underset{\sim}{R}}(u,v)\}$.

The aforementioned is a mathematical model of fuzzy snythetical evaluation. In fact, the fuzzy snythetical evaluation is using the known original image (weight matrix) and mapping (one-factor evaluation matrix) to get the result of synthetical evaluations.

In additional, we can use fuzzy integrate to form a sort of the model of the synthetical evaluation.

Suppose $U=\{u_1,u_2,\cdots,u_n\}$ is a collection consisting of n factors. $P(U)$ is U's preparing field. We have the given fuzzy vector on U,

$$M=(m_1,m_2,\cdots,m_n).$$

expresses the "chief factor". It is actually the leaders′ evaluations on the importance of n factors.

For each state on U, there is an $H\in F(U)$,

$$H=(h_1,h_2,\cdots,h_n).$$

H about "chief factor"—M′s synthetical evaluation is:

$$\int_u H(u)\circ\prod(\bullet)=H\circ M=\bigvee_{k=1}^{n}(h_k\wedge m_k)\cdot$$

The aforementioned possibility measure and fuzzy integral theory solve the rationality of the evaluation model [2].

In order to avoid the above model's defects caused by omitting the subimportant factor, we can pretreat the evaluation. For example, we can give a bottom line to every factor's satisfactory degree, when one object's satisfactory degree is below the standard, we can drive the object off the evaluation. Today, fuzzy N—integrate enlarges operator sphere of fuzzy integrate, and on the base of $(\vee,\wedge)$ some new operators appear.

We can use every sort of synthetical evaluation models resulting from different fuzzy integrate to deal with some different practical questions.

Suppose the fuzzy Vector M which stands for the possibility measures is fixed. There are m evaluators to evaluate the given object a individually, then

$$H_j=(h_{j1},h_{j2},\cdots,h_{jn})\in F(U),$$

stands for the satisfactory evaluation on α of Evaluator Number j, $j=1,2,\cdots,m$.

And to the fixed factor $u_i\in U$,

$$H_1(u_i)=h_{1i},H_2(u_i)=h_{2i},\cdots,H_m(u_i)=h_{mi}$$

can be considered as sample value whose volume is resulting from mother body U_i. U_i stands for the random variable of the objective evaluation of α on factor u_i. If you consider the social evaluation on u_i true, then the aforementioned abstraction is reasonable. From the strong law of large number in probability theory

$$p\left\{\lim_{m\to\infty}\frac{1}{m}\sum_{j=1}^{m}H_j(u_i)=h_i\right\}=1, \tag{1}$$

h_i is the mathematical expectation of the random variable U_{i1} which stands for the social evaluation on the base of social outlook on value. If we let i run from 1 to n, then we get $H\in F(U)$,

$H=(h_1,h_2,\cdots,h_n)$,

stands the social evaluation of the satisfaction on α. We can not get H directly, but we can get the sample value $H_1,H_2,\cdots,H_m$ mentioned above and then we have

$$p\left\{\lim_{m\to\infty}\frac{1}{m}\sum_{j=1}^{m}H_j=H\right\}=1. \tag{2}$$

According to the convergence theorem of fuzzy integrate series to the fuzzy integrate function H, $H_n, n=1,2,\cdots,$ on the finite field U, if

$\lim_{m\to\infty}H_n=H$,

then we can get

$$\lim_{m\to\infty}\int_U H_n(u)\circ\prod(\bullet)=\int_U H(u)\circ\prod(\bullet). \tag{3}$$

That's to say $E_0=\int_U H(u)\circ\prod(\bullet)=H_0M, E_0$ is the social synthetical evaluation to α. For the same reason, $E_j=\int_U H_j(u)\circ\prod(\bullet)=H_j\circ M$ is the synthetical evaluation of the evaluator number i on α, $j=1,2,\cdots,m.$

Combining (2) with (3), we get

$$P\left\{\lim_{m\to\infty}\int_U\frac{1}{m}\sum_{j=1}^{m}H_j(u)\circ\prod(\bullet)=\int_U H(u)\circ\prod(\bullet)\right\}=1. \tag{4}$$

In other words, when m is very big, according to probability 1

$$\int_U\frac{1}{m}\sum_{j=1}^{m}H_j(u)\circ\prod(\bullet)=E_0. \tag{5}$$

Attention, the fuzzy integral doesn't fit the common addition of function to the distributive (aw) thus, in general, we say

$$\frac{1}{m}\sum_{j=1}^{m}\int_U H_j(u)\circ\prod(\bullet)\neq\int_U\frac{1}{m}\sum_{j=1}^{m}H_j(u)\circ\prod(\bullet),$$

That is to say, we cannot use $\frac{1}{m}\sum_{j=1}^{m} E_j$ to evaluate E_0.

We call E_j the individual evaluation, $j = 1,2,\cdots,m$, and call E_0 the group evaluation true value, while $\int_U \frac{1}{m}\sum_{j=1}^{m} H_j(u) \circ \prod(\bullet)$ is called a group evaluation whose volume is m, we write it down as $\hat{E}(m)$. $\hat{E}(m)$ as an approximation of E_0, is fairer than any other E_i [3-5].

3 Application Examples

We use the fuzzy multidimensional synthetical decision and the sampling quantitative analysis to dissect the graduates' test paper in the class of 1985 and then have a comprehensive understanding of the student' standard of knowledge and capability. In this way, we can find out the factors interfering in the education quality and provide the best decision for a scientific education.

We have taken two steps to conduct the investigation:

(1) Set up an expert group, analyse the test paper and establish a fuzzy relation model.

Suppose the test paper $U = \{u_1, u_2, \cdots, u_n\}$, among these , u_i $(i=1,2,\cdots,m)$ represent test questions. Then we suppose again the investigated "knowledge" and "capability" is $V = \{v_1, v_2, \cdots, v_k, v_{k+1}, \cdots, v_n\}$, in which v_i $(i=1,2,\ldots,k)$ is the needed knowledge sample (such as mathematical analysis, higher algebra, probability theory, function of a complex variable, etc.) , v_i $(j=k+1,k+2,\ldots n)$ is the investigated capability (such as capability of operation and application). In order to have a wide investigation on the true level of students' practical "knowledge" and "capability", every question in the test paper should be included in $v_1, v_2, \cdots, v_k, v_{k+1}, \cdots, v_n$.

Now we invite some good teachers to analyse the test paper and to set up a fuzzy velation $\underset{\sim}{R} = (\underset{\sim}{r_1}, \underset{\sim}{r_2}, \cdots, \underset{\sim}{r_n}) \in F(U \times V)$ from *U* to *V*.

Suppose the "catalogue of test paper" in test paper U is:

$U = \{u_1, u_2, u_3, u_4\} =$ {filling blanks, answer questions, calculation, testimony}.

The catalogue of knowledge in test paper U is:

$V_1 = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$ = {mathematical analysis, higher algebra, analytical geometry, function of a complex variable, probability theory, higher geometry, modern algebra}.

The catalogue of capability in test paper U is:

$V_2 = \{s_1, s_2, s_3, s_4\} = \{\text{notion, operation, application, logic}\}$

Through the experts' evaluation, we can set up a fuzzy relation matrix $\underset{\sim}{R}_1$ and $\underset{\sim}{R}_2$ from U to V_1 and to V_2:

$$\underset{\sim}{R}_1 = \begin{bmatrix} 0.25 & 0.10 & 0.20 & 0.2 & 0.10 & 0 & 0.15 \\ 0.27 & 0.13 & 0.13 & 0 & 0.27 & 0.20 & 0 \\ 0.30 & 0.15 & 0.20 & 0.1 & 0.10 & 0.05 & 0.10 \\ 0.24 & 0.20 & 0.12 & 0.16 & 0 & 0.16 & 0.12 \end{bmatrix},$$

$$\underset{\sim}{R}_1 = \begin{bmatrix} 0.40 & 0.30 & 0.20 & 0.10 \\ 0.40 & 0.20 & 0.27 & 0.13 \\ 0.25 & 0.45 & 0.20 & 0.10 \\ 0.20 & 0.28 & 0 & 0.52 \end{bmatrix}.$$

In order to get the best effect of the synthetical decision, we must pay attention to science of paper U. It means that every examination question in paper U should reflect the students' "knowledge" and "capability" level correctly, at the same time the experts are needed to analyse the paper correctly and rationally.

(2) We must understand the students' standard of knowledge and ability by means of the sampling analysis of the students' paper. Only in this way can we offer the best decision to the improvement of the teaching quality and scientific administration in the future.

Let us suppose a hundred-mark system and suppose the full marks of every question are $q_1, q_2, \cdots, q_m, (\sum_{i=1}^{m} q_i = 100)$.

Now we can also suppose that the inspected object is group of $X = \{x_1, x_2, \cdots, x_a\}$ and x_i indicates a single student. The mark in question j of x_i is $a_{ij} (i = 1,2,\cdots, a, j = 1,2,\cdots, m\}$, then $A = (a_{ij})_{a\times m}$ will be a common matrix.

Let line No. j of A be divided by q_j. A will be changed into fuzzy matrix

$$\underset{\sim}{A} = (a_{ij})_{a\times m}, \text{ in which } \underset{\sim}{a}_{ij} = \frac{a_{ij}}{q_i} (i = 1,2,\cdots, a, j = 1,2,\cdots, m),$$

the preceding paragraph tells us that $\underset{\sim}{A}$ is the fuzzy relation from X to U, $i.e.$,

$\underset{\sim}{A} \in \mathscr{F}(X \times U)$ [6].

Write fuzzy transformation $\underset{\sim}{A} \circ \underset{\sim}{R} = \underset{\sim}{B} = (b_{ij})_{a\times n}$, and $\underset{\sim}{B}$ is the fuzzy relation from X to V: $\underset{\sim}{B} = \mathscr{F}(X \times V)$, in it $b_{i1}, b_{i2}, \cdots, b_{in}$, b_{in} indicate the result of the inspected x_i, that is , the real knowledge and ability level of x_i.

We must see that if Line No j of R is all nought. Line No j of B is also nought no matter what A is. Thus we can't investigate the condition about V_i of the students from $\underset{\sim}{B}$. So a good set of paper U needs $\underset{\sim}{B}$ to satisfy

$$\underset{i=1}{\vee} x_{ij} = 1.$$

To sum up, we can sample and inspect five pieces of paper from the students in the Class of 1985. Suppose $X = \{x_1, x_2, x_3, x_4, x_5\}$, and x_i indicates a single student. The fuzzy relation matrix of the mark that the single student i makes in Type j is:

$$A = \begin{bmatrix} 0.90 & 0.80 & 0.88 & 0.80 \\ 0.80 & 0.53 & 0.73 & 0.60 \\ 0.60 & 0.87 & 0.50 & 0.48 \\ 0.80 & 0.80 & 0.63 & 0.40 \\ 0.85 & 0.67 & 0.80 & 0.44 \end{bmatrix}.$$

Normalize $\underset{\sim}{A}$, and we will get:

$$A' = \begin{bmatrix} 0.26 & 0.24 & 0.26 & 0.24 \\ 0.30 & 0.20 & 0.27 & 0.23 \\ 0.24 & 0.36 & 0.20 & 0.20 \\ 0.30 & 0.30 & 0.24 & 0.16 \\ 0.31 & 0.24 & 0.29 & 0.16 \end{bmatrix}$$

and

$$\underset{\sim}{B_1} = \underset{\sim}{A'} \circ \underset{\sim}{R_1} = \begin{bmatrix} 0.26 & 0.20 & 0.20 & 0.20 & 0.24 & 0.20 & 0.15 \\ 0.27 & 0.20 & 0.20 & 0.20 & 0.20 & 0.20 & 0.15 \\ 0.27 & 0.20 & 0.20 & 0.20 & 0.27 & 0.20 & 0.15 \\ 0.27 & 0.16 & 0.20 & 0.20 & 0.27 & 0.20 & 0.15 \\ 0.29 & 0.15 & 0.20 & 0.20 & 0.24 & 0.20 & 0.15 \end{bmatrix},$$

$$\underset{\sim}{B}_2 = \underset{\sim}{A}' \circ \underset{\sim}{R}_2 = \begin{bmatrix} 0.26 & 0.26 & 0.24 & 0.24 \\ 0.30 & 0.30 & 0.20 & 0.23 \\ 0.36 & 0.24 & 0.27 & 0.20 \\ 0.30 & 0.30 & 0.27 & 0.16 \\ 0.31 & 0.30 & 0.24 & 0.16 \end{bmatrix},$$

thus we can investigate the knowledge and ability level of the students x_1, x_2, x_3, x_4, x_5.

For example, in order to inspect the "ability" of student x_3, in $\underset{\sim}{B}_2$

$$\max\{0.36, 0.24, 0.27, 0.20\} = 0.36,$$
$$\min\{0.36, 0.24, 0.27, 0.20\} = 0.20,$$

We can see that student x_3 grasps the fundamental concept well, but his logic inference ability is quite poor.

The average value of every element of $\underset{\sim}{B}_2$ is :

$$(0.36, 0.28, 0.244, 0.198).$$

From above we can see that the student grasps fundamental concept well, but his logical inference ability is poor. So we must give special attention to he training of the logic inference ability in our future teaching.

The job assignment of the college is rather complicated, which touches upon a lot of objects and factors. How to use the talented persons rationally is the crux in the job assignment. So the author has applied the fuzzy multidimensional synthetical decision method to change the qualitative analysis of the job assignment of the graduates into quantative analysis, which shows a better mathematical model of the job assignment and offers a reliable scientific basis.

The fuzzy multidimensional synthetical decision can be widely used in every field in natural science and social science [7]. For example, the consultative system of the medical experts and the computer interrogation, which marks the modernization in medical diagnosis, is an excellent achievement in scientific research that combines the fuzzy multidimensional synthetical decision with practice. The fundamental procedure in diagnosis is:

(symptom) (patient) $\boxed{\underset{\sim}{A}}$ import

(diagnosis) →(symptom) $\boxed{\underset{\sim}{R}}$ experts system

(diagnosis) →(patient) $\boxed{\underset{\sim}{B}}$ export

The main model:

The experts' experience in treating or curing the patients →mathematicism→computer study→feedback revision →consultative system of experts →computer interrogation.

It is not difficult to realize "the sensor interrogating" with the help of the fuzzy transformation and computers. In fact, in China, the fuzzy mathematical model that Doctor Guang Youbo uses to cure liver diseases has been made into the software of the expert system, and the computer interrogation has been realized successfully.

4 Conclusion

The best fuzzy multidimensional synthetical decision is an optimization. This method is often used in solving the problem with multiple targets and factors, which is difficult to evaluate, but it can be settled by the fuzzy transformation [8]. If we can program the computer and realize the automation of the evaluation, the effect will be better. In addition, if we can choose the proper factors $u_1, u_2, \ldots u_n$ and show the weighted numbers of every factor, and if those who take part in evaluation possess representative and practical experience, there is a great significance to improve the effect of the fuzzy multidimensional synthetical decision.

References

1. Zimmermann, H.J.: Fuzzy sets theory and its applications. Kluwer Nijhoff Pubilishing, Boston (1985)
2. Saaty, T.: Exploring the interface between Hierachies multiple objectives and fuzzy Sets. Fuzzy Sets and Systems (1), 57–68 (1978)
3. Rundell, W.: Proceedings Amer. Math. Soc., vol. 4, pp. 190–223 (1987)
4. Qui, B.H.: Weighted Hardy Spaces. Math Nachr. (103), 45–62
5. Bellman, R.E., Zadeh, L.A.: Decision—Making in a Fuzzy Environment. Manage Sci. 1970(17), 141–164 (1981)
6. Jones, A., Kaufmann, A., Zimmermann, H.J.: Fuzzy sets theory and applications. D. Reidel Publishing Company, Dordrecht (1986)
7. Zadeh, L.A.: Fuzzy sets. Inform. Contr1 (8), 338–353 (1965)
8. Blin, J.M.: Fuzzy relations in group decision theory. Journal Cyber. (4), 17–22 (1974)

$\boldsymbol{\omega\theta}$-Countability in an $\boldsymbol{L\omega}$-Space

Zhao-xia Huang and Shui-li Chen

School of Science, Jimei University, Xiamen, Fujian, 361021, China
e-mail: {zxhuang,sgzx}@jmu.edu.cn

Abstract. An $L\omega$ -space is view as a fuzzy topology space containing various closure operators. In this paper, some new notions of the first and the second $\omega\theta$ -countability are introduced in an $L\omega$ -space. Some basic properties of them are respectively given. For example, the first and the second $\omega\theta$-countability are both hereditary property, countable multiplicative property and invariant property under $(\omega_1, \omega_2)\theta$ -homomorphism.

Keywords: $L\omega$-space, $\omega\theta$ -open set, $\omega\theta R$-neighborhood base, the first $\omega\theta$-countable space, the second $\omega\theta$-countable space, $\omega\theta$-base.

1 Introduction

The countability theory is one of the most important theories in topology. In 1988, Wang introduced the concept of R-neighborhood and established the countability theory in an L-fuzzy topology space [1]. In 2002, Chen and Dong further generalized the above notions and established an L-fuzzy order-preserving operator space [2], then the ω-countability [3], the ω-connectedness [4] and the ω-separation [5] were given respectively. In this paper, we will enrich the countability theory. We will present two new countable spaces which is called the first $\omega\theta$-countable space and the second $\omega\theta$-countable space and discuss their properties.

2 Preliminary Concepts and Notations

Throughout this paper, L denotes a fuzzy lattice, i.e., a completely distributive lattice with order-reversing involution "$\prime$". An element a of L is said to be $\vee$-irreducible (or a molecule) [1] if $a \leq b \vee c$ implies that $a \leq b$ or $a \leq c$, where $b, c \in L$. The set consisting of all nonzero $\vee$-irreducible elements of L will be denoted by M, and the greatest element and the least element of L will be denoted by 1 and 0, respectively. For each non-empty crisp set X, L^X

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 519–526.
springerlink.com

denotes all L-fuzzy sets on X and with value in L, $M^*(L^X)$ denotes the set of all molecules, i.e., nonzero $\vee$-irreducible L-fuzzy points [6] in L^X, and the constant L-fuzzy set taking on the constant values 1 and 0 at each x in X will be denoted by 1_X and 0_X, respectively.

Definition 1. *[2,4] Let X be a non-empty crisp set.*

(a)An operator $\omega : L^X \to L^X$ is called an ω-operator if (1) $\omega(1_X) = 1_X$; (2) For each $A, B \in L^X$, if $A \leq B$, then $\omega(A) \leq \omega(B)$; (3) For each $A \in L^X$, $A \leq \omega(A)$.

(b) An L-set $A \in L^X$ is called an ω-set if $A = \omega(A)$.

(c) Put $\Omega = \{A \in L^X \mid A = \omega(A)\}$, and we will call the pair (L^X, Ω) an L-fuzzy order-preserving operator space,or an $L\omega$-space.

Definition 2. *[2,4] Let (L^X, Ω) be an $L\omega$ -space, $P \in L^X$ and $x_\alpha \in M^*(L^X)$. If there exists a $Q \in \Omega$, such that $x_\alpha \not\leq Q$ and $P \leq Q$, we will call P an ωR-neighborhood of x_α . We will denote by $\omega\eta(x_\alpha)$ the collection of all ωR-neighborhoods of x_α*

Definition 3. *[7] Let (L^X, Ω) be an $L\omega$-space, $A \in L^X$, and $x_\alpha \in M^*(L^X)$. If $A \not\leq P$ for each $P \in \omega\eta(x_\alpha)$, then we will call x_α an ω-adherence point of A. We will call the union of all ω-adherence points of A the ω-closure of A, and will denote by $\omega cl(A)$. If $A = \omega cl(A)$, then we will call A an ω-closed set. If A is an ω-closed set, then we will call A' an ω-open set. If $Q = \omega cl(Q)$ and $x_\alpha \not\leq Q$, then Q is said to be an ω-closed R-neighborhood (briefly, ωCR-neighborhood) of x_α. We will denote by $\omega\eta^-(x_\alpha)$ the collection of all ωCR-neighborhoods of x_α.*

Definition 4. *[2,4] Let (L^X, Ω) be an $L\omega$-space, $A \in L^X$. Put $\omega int(A) = \vee\{B \in L^X | B \leq A$ and B is an ω-open set in $L^X\}$. We will call $\omega int(A)$ the ω-interior of A. Obviously, A is an ω-open set if and only if $A = \omega int(A)$.*

Definition 5. *Let (L^X, Ω) be an $L\omega$-space, $P \in L^X$ and $x_\alpha \in M^*(L^X)$. If there exists a $Q \in \Omega$ such that $x_\alpha \not\leq \omega int(Q)$ and $P \leq Q$,we will call P an $\omega\theta R$-neighborhood of x_α. We will denote by $\omega\theta\eta(x_\alpha)$ the collection of all $\omega\theta R$-neighborhood of x_α. If $A \not\leq \omega int(P)$ for each $P \in \omega\eta(x_\alpha)$, then we will call x_α an $\omega\theta$-adherence point of A. We will call the union of all $\omega\theta$-adherence points of A the $\omega\theta$-closure of A, and will denote by $\omega\theta cl(A)$. If $A = \omega\theta cl(A)$, then we will call A an $\omega\theta$-closed set. If A is an $\omega\theta$-closed set, then we will call A' an $\omega\theta$-open set. Let $\omega\theta o(L^X)$ be the union of all $\omega\theta$-open sets. If $Q \in (L^X)$ is an $\omega\theta$-closed set and $x_\alpha \not\leq \omega int(Q)$, then Q is said to be an $\omega\theta$-closed R-neighborhood (briefly, $\omega\theta CR$-neighborhood) of x_α. We will denote by $\omega\theta\eta^-(x_\alpha)$ the collection of all $\omega\theta CR$-neighborhood of x_α.*

It is obvious that an ω-adherence point of A is an $\omega\theta$-adherence point of A, an ωR-neighborhood (ω-closed R-neighborhood) of $x_\alpha$is an $\omega\theta R$-neighborhood ($\omega\theta$-closed R-neighborhood) of x_α.

Definition 6. *Let (L^X, Ω) be an $L\omega$-space, $e \in M^*(L^X)$, $\beta \subset \omega\theta o(L^X)$, $\mu \subset \omega\theta\eta^-(e)$.*

1) If for each $G \in \omega\theta o(L^X)$, there exists a subfamily φ of β such that $G = \vee\{B|B \in \varphi\}$, then β is called an $\omega\theta$-base of (L^X, Ω);

2) For each $\gamma \subset \omega\theta o(L^X)$, if the family of all intersection of finite elements in γ is an $\omega\theta$-base of (L^X, Ω), then γ is called an $\omega\theta$-subbase of (L^X, Ω);

3) If for each $P \in \omega\theta\eta^-(e)$, there exists a $Q \in \mu$ such that$P \leq Q$, then μ is called an $\omega\theta R$-neighborhood-base of e.

3 The Second $\omega\theta$-Countable Space

In this section, the second $\omega\theta$-countable space and some properties are introduced, including hereditary property, countable multiplicative property and invariant property under $(\omega_1, \omega_2)\theta$-homomorphism.

Definition 7. *Let (L^X, Ω) be an $L\omega$-space. (L^X, Ω) is called the second $\omega\theta$-countability space, briefly, $\omega\theta C_2$-space, if there is a countable $\omega\theta$-base in (L^X, Ω).*

Theorem 1. *Let $(L_i^{X_i}, \Omega_i)(i = 1, 2)$ be two $L\omega$-spaces and $f : (L_1^{X_1}, \Omega_1) \to (L_2^{X_2}, \Omega_2)$ be $(\omega_1, \omega_2)\theta$-continuous and $(\omega_1, \omega_2)\theta$-open surjective order homomorphism. If $(L_1^{X_1}, \Omega_1)$ is an $\omega_1\theta C_2$-space, then $(L_2^{X_2}, \Omega_2)$ is an $\omega_2\theta C_2$-space.*

Proof. Suppose that B is an $\omega_2\theta$-open set in $(L_2^{X_2}, \Omega_2)$. Then $f^{-1}(B)$ is an $\omega_1\theta$-open set in $(L_1^{X_1}, \Omega_1)$ according to the $(\omega_1, \omega_2)\theta$-continuity of f . Since $(L_1^{X_1}, \Omega_1)$ is an $\omega_1\theta C_2$-space, there exists a countable $\omega_1\theta$-base β in $(L_1^{X_1}, \Omega_1)$, such that $f^{-1}(B)$ can be represented by the union of some members of β, that is, there exists a subfamily φ of β such that $f^{-1}(B) = \vee\{A|A \in \varphi\}$. Then $B = \vee\{f(A)|A \in \varphi\}$ holds because f is surjective. However f is $(\omega_1, \omega_2)\theta$-open, then $f(A)$is an $\omega_2\theta$-open set in $(L_2^{X_2}, \Omega_2)$. Put $\gamma = \{f(A)|A \in \beta\}$, then γ is an $\omega_2\theta$-base in $(L_2^{X_2}, \Omega_2)$. Since β is countable, γ is also countable. Hence, γ is the countable base in $(L_2^{X_2}, \Omega_2)$, that is, $(L_2^{X_2}, \Omega_2)$ is an $\omega_2\theta C_2$-space.

Preceding theorem shows that the second $\omega\theta$-countability is homomorphism.

Definition 8. *[3] Let (L^X, Ω) be an $L\omega$-space, Y be a nonempty subset of X and $\Omega|_Y = \{A|_Y|A \in \Omega\}$, where $A|_Y$ is the restriction of A on Y, i.e., for each $y \in Y, A|_Y(y) = A(y)$, then $(L^Y, \Omega|_Y)$ is called an ω-subspace of (L^X, Ω).*

Theorem 2. *Let $(L^Y, \Omega|_Y)$ be an ω-subspace of $L\omega$-space (L^X, Ω). Then the following statements hold.*

1) If G is an $\omega\theta$-closed set of (L^X, Ω), then $G|_Y$ is an $\omega\theta$-closed set of $(L^Y, \Omega|_Y)$;

2) If H is an $\omega\theta$-closed set of $(L^Y, \Omega|_Y)$, then there exists an $\omega\theta$-closed set G of (L^X, Ω), such that $G|_Y = H$;

3) If E is an $\omega\theta$-open set of (L^X, Ω), then $E|_Y$ is an $\omega\theta$-open set of $(L^Y, \Omega|_Y)$;

4) If F is an $\omega\theta$-open set of $(L^Y, \Omega|_Y)$, then there exists an $\omega\theta$-open set E of (L^X, Ω), such that $E|_Y = F$.

Proof. 1) Let G be an $\omega\theta$-closed set of (L^X, Ω), $e \in M^*(L^Y)$ and $e \leq \omega\theta cl(G|_Y)$, then for each $P \in \omega\theta\eta(e)$, there exists a $Q \in \Omega$, such that $P \leq Q|_Y, e \not\leq wint(Q|_Y)$, and $G|_Y \not\leq \omega int(Q|_Y)$. According to Theorem 1 on [8], we have $G \not\leq \omega int(Q), e^* \not\leq \omega int(Q)$ and $P^* \leq Q$, i.e.,$P^* \in \omega\theta\eta(e^*)$, and $e^* \leq \omega\theta cl(G)$, where e^* and P^* stand for the extension of e and P in X respectively. Because G is an $\omega\theta$-closed set, we get $e^* \leq G$ and $e \leq G|_Y$, that is, $G|_Y$ is an $\omega\theta$-closed set of $(L^Y, \Omega|_Y)$.

2)Let H be an $\omega\theta$-closed set of $(L^Y, \Omega|_Y)$, then $H = (\omega\theta cl(H^*))|_Y$. In fact, if $e \leq H$, then $e^* \leq H^* \leq \omega\theta cl(H^*)$, hence $e \leq (\omega\theta cl(H^*))|_Y$. Conversely, if $e \leq (\omega\theta cl(H^*))|_Y$, then $e^* \leq \omega\theta cl(H^*)$, and for each $P \leq \omega\theta\eta(e)$, $P^* \leq \omega\theta\eta(e^*)$, hence $H^* \not\leq \omega int(P^*)$. According to Theorem 1 on [8], we have $H \not\leq \omega int(P)$, hence $e \leq \omega\theta cl(H) = H$, $H = (\omega\theta cl(H^*))|_Y$. According to Theorem 2.1 on [2], we know that $\omega\theta cl(H^*)$ is an $\omega\theta$ -closed set of (L^X, Ω).

For each $E \in L^X, E'|_Y = (E|_Y)'$, we know that 3) and 4) are true by 1) and 2) respectively.

Theorem 3. *Let $(L^Y, \Omega|_Y)$ be an ω-subspace of $L\omega$-space (L^X, Ω). Then the following statements hold.*

1) If β is an $\omega\theta$-base of (L^X, Ω), then $\beta|_Y$ is an $\omega\theta$-base of $(L^Y, \Omega|_Y)$;

2) If γ is an $\omega\theta$-subbase of (L^X, Ω), then $\gamma|_Y$ is an $\omega\theta$-subbase of $(L^Y, \Omega|_Y)$.

Proof. 1) Let H be an $\omega\theta$-open set of $(L^Y, \Omega|_Y)$, then according to Theorem 3.2, there exists an $\omega\theta$-open set G of (L^X, Ω), such that $G|_Y = H$. Since β is an $\omega\theta$-base of (L^X, Ω), there exists a subfamily φ of β such that $G = \vee\{B|B \in \varphi\}$. According to Theorem 2.7.2 on [1], $G|_Y = (\vee\{B|B \in \varphi\})|_Y = \vee\{B|_Y|B \in \varphi\}$. Since $\varphi|_Y = \{B|_Y|B \in \varphi\} \subset \beta|_Y$, $\beta|_Y$ is an $\omega\theta$-base of $(L^Y, \Omega|_Y)$.

2) Let β_0 be the union of all intersection of finite elements in $\gamma|_Y$, i.e., $\beta_0 = \wedge\{S_t|_Y|t \in T, S_t \in \gamma, T$ is a finite index set$\}$, then $\wedge\{S_t|_Y|t \in T\} = (\wedge\{S_t|t \in T\})|_Y$ [1]. Let $\beta = \wedge\{S_t|t \in T, S_t \in \gamma, T$ is an finite index set $\}$. Since γ is an $\omega\theta$-subbase of (L^X, Ω), β is an $\omega\theta$-base of (L^X, Ω) and $\beta_0 = \beta|_Y$, hence β_0 is an $\omega\theta$-base of $(L^Y, \Omega|_Y)$ and $\gamma|_Y$ is an $\omega\theta$-subbase of $(L^Y, \Omega|_Y)$.

Theorem 4. *Let $(L^Y, \Omega|_Y)$ be an ω-subspace of $L\omega$-space (L^X, Ω). If (L^X, Ω) is an $\omega\theta C_2$-space, then $(L^Y, \Omega|_Y)$ is also an $\omega\theta C_2$-space.*

Proof. Let β be the countable $\omega\theta$-base of (L^X, Ω), then $\beta|_Y$ is the countable $\omega\theta$-base of $(L^Y, \Omega|_Y)$ according to Theorem 3.3, that is to say, $(L^Y, \Omega|_Y)$ is also an $\omega\theta C_2$-space.

Preceding theorem shows that the second $\omega\theta$-countability is hereditary.

Theorem 5. *Suppose that $\{(L^{X_t}, \Omega_t)|t \in T\}$ is a class of a countable number of $L\omega_t$-space and (L^X, Ω) is their product space. Then the following results hold.*

1) If $\forall t \in T, (L^{X_t}, \Omega_t)$ *is an* $\omega_t\theta C_2$*-space, then* (L^X, Ω) *is an* $\omega\theta C_2$*-space;*

2) If (L^X, Ω) *is an* $\omega\theta C_2$*-space, and* $\forall t \in T, (L^{X_t}, \Omega_t)$ *is a stratified space, then* $\forall t \in T, (L^{X_t}, \Omega_t)$ *is an* $\omega_t\theta C_2$*-space.*

Proof. 1) Put $\beta = \{\prod B_t | B_t \in \beta_t, t \in T\}$, where β_t is the countable $\omega_t\theta$-base of (L^{X_t}, Ω_t). According to the countability of T and β_t, we know that β is the countable $\omega\theta$-base of (L^X, Ω). Hence, (L^X, Ω) is an $\omega_t\theta C_2$-space.

2)Suppose that (L^X, Ω) is an $\omega\theta C_2$-space and $\forall t \in T, (L^{X_t}, \Omega_t)$ a stratified space. Obviously, the projective mapping $P_t : L^X \to L^{X_t}$ is $(\omega, \omega_t)\theta$-continuous and $(\omega, \omega_t)\theta$-open surjective order homomorphism. According to Theorem 3.1,(L^{X_t}, Ω_t) is an $\omega_t\theta C_2$-space.

Preceding theorem shows that the second $\omega\theta$-countability is countable multiplicative.

4 The First $\omega\theta$-Countable Space

In this section, the first $\omega\theta$-countable space and some properties are introduced, including hereditary property, countable multiplicative property and invariant property under $(\omega_1, \omega_2)\theta$-homomorphism.

Definition 9. *Let* (L^X, Ω) *be an* $L\omega$*-space. If for any* $e \in M^*(L^X)$*, there exists a countable* $\omega\theta R$*-neighborhood base* $\mu(e)$*, then we call* (L^X, Ω) *the first* $\omega\theta$*- countable space, briefly,* $\omega\theta C_1$*-space.*

Theorem 6. $\omega\theta C_2$*-space must be* $\omega\theta C_1$*-space.*

Proof. Suppose that (L^X, Ω) is an $\omega\theta C_2$-space, then there exists a countable $\omega\theta$-base β. For any $e \in M^*(L^X)$, put $\mu(e) = \{Q \in \beta' | e \not\leq Q\}$, where Q is the $\omega\theta CR$-neighborhood of e. Then $\mu(e)$ is a countable $\omega\theta R$-neighborhood base of e. Hence (L^X, Ω) is an $\omega\theta C_1$-space.

Theorem 7. *Let* $(L_i^{X_i}, \Omega_i)(i = 1, 2)$ *be two* $L\omega$*-spaces and* $f : (L_1^{X_1}, \Omega_1) \to (L_2^{X_2}, \Omega_2)$ *be* $(\omega_1, \omega_2)\theta$*-continuous and closed one-to-one surjective order homomorphism. If* $(L_1^{X_1}, \Omega_1)$ *is an* $\omega_1\theta C_1$*-space, then* $(L_2^{X_2}, \Omega_2)$ *is an* $\omega_2\theta C_1$*-space.*

Proof. Suppose that $d \in M^*(L_2^{X_2})$, then there exists a $e \in M^*(L_1^{X_1})$, such that $f(e) = d$. Let $Q \in \omega_2\theta\eta^-(d)$. According to the $(\omega_1, \omega_2)\theta$-continuity of f, we have $f^{-1}(Q) \in \omega_1\theta\eta^-(e)$. Since $(L_1^{X_1}, \Omega_1)$ is an $\omega_1\theta C_1$-space, there exists a countable $\omega_1\theta R$- neighborhood base $\mu(e)$ of e. Then there exists a $P \in \mu(e)$, such that $f^{-1}(Q) \leq P$, i.e.,$Q \leq f(P)$. Hence from $e \not\leq P$, we have $f(e) \not\leq f(P)$. Since f is $(\omega_1, \omega_2)\theta$-closed order homomorphism, $f(P) \in \omega_2\theta\eta^-(d)$. Put $\nu(d) = \{f(P)|P \in \mu(e)\}$, then $\nu(d)$ is a countable $\omega_2\theta R$-neighborhood base of d. Hence, $(L_2^{X_2}, \Omega_2)$ is an $\omega_2\theta C_1$-space. Preceding theorem shows that the first $\omega\theta$-countability is homomorphism.

Theorem 8. *Let* $(L^Y, \Omega|_Y)$ *be an* ω*-subspace of* $L\omega$*-space* (L^X, Ω)*. If* (L^X, Ω) *is an* $\omega\theta C_1$*-space, then* $(L^Y, \Omega|_Y)$ *is also an* $\omega\theta C_1$*-space.*

Proof. Put $e \in M^*(L^Y), P \in \omega\theta\eta^-(e)$. According to Theorem 3.2, we have $P \in \omega\theta\eta^-(e^*)|_Y$, i.e., there exists a $Q \in \omega\theta\eta^-(e^*)$, such that $P = Q|_Y$. Since (L^X, Ω) is an $\omega\theta C_1$-space, there exists a countable $\omega\theta R$-neighborhood base $\mu(e^*)$ of e^*, such that $G \in \mu(e^*)$ and $Q \leq G$. Put $\nu(e) = \mu(e^*)|_Y$, then we have $G|_Y \in \nu(e)$ such that $P \leq G|_Y$, hence $\nu(e)$ is a countable $\omega\theta R$-neighborhood base of e, that is, $(L^Y, \Omega|_Y)$ is also an $\omega\theta C_1$-space.

Preceding theorem shows that the first $\omega\theta$-countability is hereditary.

Theorem 9. *Suppose that (L^X, Ω) is an $\omega\theta C_1$-space, then for any $e \in M^*(L^X)$, there exists a countable $\omega\theta R$-neighborhood base of e, $\mu^*(e) = \{P_1, P_2, \cdots\}$, satisfying that $P_1 \leq P_2 \leq \cdots$.*

Proof. Since (L^X, Ω) is an $\omega\theta C_1$-space, for any $e \in M^*(L^X)$, there exists a countable $\omega\theta R$-neighborhood base of e, $\mu(e) = \{Q_1, Q_2, \cdots\}$. Put $P_n = Q_1 \vee Q_2 \vee \cdots Q_n (n = 1, 2, \cdots)$, then $\forall G \in \omega\theta\eta^-(e)$, there exists a $Q_t \in \mu(e)$, such that $G \leq Q_t$, hence there exists a n, such that $G \leq P_n$. For any $Q_i \in \mu(e), e \not\leq \omega int(Q_i)$. Since $P_n = Q_1 \vee Q_2 \vee \cdots Q_n (n = 1, 2, \cdots)$ are $\omega\theta$-closed sets, $e \not\leq \omega int(P_n), P_n \in \omega\theta\eta^-(e)(n = 1, 2, \cdots)$. Hence $\mu^*(e) = \{P_1, P_2, \cdots\}$ is $\omega\theta R$-neighborhood base of e, satisfying that $P_1 \leq P_2 \leq \cdots$.

Definition 10. *[9] Let (L^X, Ω) be an $L\omega$-space, $x_\alpha \in M^*(L^X)$ and $N = \{N(n) \in M^*(L^X)|n \in D\}$ is a molecular net in L^X. If $\forall P \in \omega\eta^-(x_\alpha)$, there exists a $m \in D$, such that $N(n) \not\leq \omega int(P)$ whenever $n \geq m$. i.e., $N(n)$ is not in P eventually, then x_α is said to be an $\omega\theta$-limit point of N, or called that $N\omega\theta$-converges to x_α.*

Theorem 10. *Suppose that (L^X, Ω) is an $\omega\theta C_1$-space, $A \in L^X$ and $e \in M^*(L^X)$, then $e \leq \omega\theta cl(A)$ if and only if e is the $\omega\theta$-limit point of some molecular sequence $S = \{S(n)|n \in N\}$ of A.*

Proof. Let $e \leq \omega\theta cl(A)$. According to Theorem 4.4, there exists a countable $\omega\theta R$-neighborhood base of e, $\mu^*(e) = \{P_1, P_2, \cdots\}$, satisfying that $P_1 \leq P_2 \leq \cdots$, and $\forall n \in N, A \not\leq \omega int(P_n)$. Hence there exists a molecule $S(n) \leq A$, such that $S(n) \not\leq \omega int(P_n)$. Put $S = \{S(n)|n \in N\}$, then $\forall Q \in \omega\theta\eta^-(e)$, there exists a $P_n \in \mu^*(e)$, such that $Q \leq P_n$ and $S(n) \not\leq \omega int(Q)$, hence e is an $\omega\theta$-limit point of S.

Conversely, let $S = \{S(n)|n \in N\}$ be a molecular sequence of A which $\omega\theta$-limit point is e, then $\forall Q \in \omega\theta\eta^-(e)$, there exists a $m \in N$, such that $S(n) \not\leq \omega int(Q)$ whenever $n \geq m$, hence $A \not\leq \omega int(Q)$. This means that e is an $\omega\theta$-adherence point of A, i.e., $e \leq \omega\theta cl(A)$.

Theorem 11. *Suppose that $(L_1^{X_1}, \Omega_1)$ is an $\omega_1\theta C_1$-space, $(L_2^{X_2}, \Omega_2)$ is any $L\omega_2$-space, and $f : (L_1^{X_1}, \Omega_1) \to (L_2^{X_2}, \Omega_2)$ is any order homomorphism. Then f is $(\omega_1, \omega_2)\theta$-continuous at a molecule $e \in M^*(L_1^{X_1})$ if and only if $f(S)$ is the molecular sequence that $\omega_2\theta$-converges to $f(e)$ in $L_2^{X_2}$, whenever S is the molecular sequence that $\omega_1\theta$-converges to e in $L_1^{X_1}$.*

Proof. Suppose that f is $(\omega_1, \omega_2)\theta$-continuous at a molecule $e \in M^*(L_1^{X_1})$, $S = \{S(n)|n \in N\}$ is the molecular sequence that $\omega_1\theta$-converges to e in $L_1^{X_1}$, $Q \in \omega_2\theta\eta^-(f(e))$, then $f^{-1}(Q) \in \omega_1\theta\eta^-(e)$. Hence there exists a $m \in N$, such that $S(n) \not\leq \omega int(f^{-1}(Q))$ whenever $n \geq m$, i.e., $f(S(n)) \not\leq \omega int(Q)$. That is to say, $f(S) = \{f(S(n))|n \in N\}$ is the molecular sequence that $\omega_2\theta$-converges to $f(e)$ in $L_2^{X_2}$.

Conversely, if f is not $(\omega_1, \omega_2)\theta$-continuous at $e \in M^*(L_1^{X_1})$, then there exists a $Q \in \omega_2\theta\eta^-(f(e))$, such that $f^{-1}(Q) \notin \omega_1\theta\eta^-(e)$. Since $(L_1^{X_1}, \Omega_1)$ is an $\omega_1\theta C_1$-space, there exists a countable $\omega_1\theta R$- neighborhood base of e, $\mu^*(e) = \{P_1, P_2, \cdots\}$, satisfying that $P_1 \leq P_2 \leq \cdots$ and $f^{-1}(Q) \not\leq \omega int(P_n)$. Hence there exists a molecule $S(n) \leq f^{-1}(Q)$ and $S(n) \not\leq \omega int(P_n)$. Put $S = \{S(n)|n \in N\}$, then S is the molecular sequence that $\omega_1\theta$-converges to e in $f^{-1}(Q)$. Notice that $S(n) \leq f^{-1}(Q)$, we have $f(S(n)) \leq Q$, this means that $f(e)$ is not the $\omega_2\theta$-limit point of the molecular sequence $f(S) = \{f(S(n))|n \in N\}$.

Theorem 12. *Suppose that $\{(L^{X_t}, \Omega_t)|t \in T\}$ is a class of a countable number of $\omega_t\theta C_1$-space and (L^X, Ω) is their product space, then (L^X, Ω) is also an $\omega\theta C_1$-space.*

Proof. Let $x = \{x_t\}_{t\in T}$ is any point of $X, \alpha \in M$, then $x_\alpha \in M^*(L^X)$ and $\forall t \in T, (x_t)_\alpha \in M^*(L^{X_t})$. Since $\forall t \in T, (L^{X_t}, \Omega_t)$ is an $\omega_t\theta C_1$-space, for any $(x_t)_\alpha$, there exists a countable $\omega_t\theta R$-neighborhood base μ_t. Put $\mu = \{\vee(P_t^{-1}(A_t))|A_t \in \mu_t, t \in S, S$ is a finite subset of $T\}$, where $P_t : L^X \to L^{X_t}$ is a projective order homomorphism mapping. According to the countability of T, we know that μ is countable in L^X. Now we prove that μ is an $\omega\theta R$-neighborhood base of x_α.

In fact, by the continuity of P_t, we know that $\forall t \in T, (P_t^{-1}(A_t)) \in \omega\theta cl(L^X)$. Since μ_t is an $\omega_t\theta R$- neighborhood base of $(x_t)_\alpha$ and $A_t \in \mu_t, \alpha \not\leq \omega int(A_t(x_t))$, $P_t^{-1}(A_t(x)) = A_t(P_t(x)) = A_t(x_t)$, i.e. $\alpha \not\leq \omega int(P_t^{-1}(A_t(x)))$, or $\forall t \in S, (P_t^{-1}(A_t)) \in \omega\theta\eta^-(x_\alpha)$. Notice that S is a finite set $(\vee(P_t^{-1}(A_t))|t \in S) \in \omega\theta\eta^-(x_\alpha)$. Let Q be an $\omega\theta CR$-neighborhood of x_α, then according to the Theorem 16 on [10] and $\beta = \{\wedge(P_t^{-1}(S_t))|S_t \in \omega\theta o(L^{X_t}), t \in F, F$ is a finite subset of $T\}$ is an $\omega\theta$-base of (L^X, Ω), we know that x_α is an $\omega\theta CR$-neighborhood formed by $H = \vee\{(P_k^{-1}(B_k))|k \in K\}$, such that $H \geq Q$, where K is a finite subset of T and $\forall k \in K, B_k$ is an $\omega_t\theta$-closed set of (L^{X_t}, Ω_t). Since H is an $\omega\theta CR$-neighborhood of x_α, we know that B_k is an $\omega_t\theta CR$-neighborhood of $(x_t)_\alpha$. Because μ_k is an $\omega_k\theta CR$-neighborhood of $(x_k)_\alpha$, there exists a $A_k \in \mu_k$ such that $A_k \geq B_k$, hence $\vee\{(P_k^{-1}(A_k))|k \in K\} \in \mu$ and $\vee\{(P_k^{-1}(A_k))|k \in K\} \geq H \geq Q$. This means that μ is an $\omega\theta R$-neighborhood base of x_α. Hence, (L^X, Ω) is also an $\omega\theta C_1$-space.

Preceding theorem shows that the first $\omega\theta$-countability is countable multiplicative.

5 Conclusion

In this paper, starting with the concepts which are called an ω-operator and an $L\omega$-space, we introduce the concepts of $\omega\theta$-open set, $\omega\theta R$-neighborhood base, the first $\omega\theta$-countable space, the second $\omega\theta$-countable space, discuss their basic properties, such as the first and the second $\omega\theta$- countable space are both hereditary property, countable multiplicative property and invariant property under $(\omega_1, \omega_2)\theta$-homomorphism. All the discussions will offer a theoretical foundation in fuzzy operator.

Acknowledgements. The project is supported by the Natural Science Foundation of Fujian Province, China(NSFC No.A0810014).

References

1. Wang, G.J.: L-fuzzy Topology Space Theory. Shanxi Normal University Press, Xi'an (1988)
2. Chen, S.L., Dong, C.Q.: L-fuzzy Order-preserving Operator Spaces. Fuzzy System and Mathematics 16(special issue), 36–41 (2002)
3. Chen, S.L.: ω-countability on L-order-preserving Operator Spaces. Fuzzy System and Mathematics 18(3), 11–16 (2004)
4. Huang, Z.X.: The ω-connected Properties on an LF Order-preserving Operator Space. Mathematics Magazine 27(3), 343–347 (2007)
5. Huang, Z.X., Chen, S.L.: The ω-separations on L-fuzzy Order-preserving Operator Spaces. Mathematics Magazine 25(4), 383–388 (2005)
6. Liu, Y.M., Luo, M.K.: Fuzzy Topology. In: Advances in Fuzzy Systems Applications and Theory, vol. 9. World Scientific Publishing Co. Pte.Ltd., Singapore (1997)
7. Chen, S.L., Chen, G.L., Zhang, J.C.: ω-Convergence Theory of Filters in Lω-spaces. Fuzzy Optim Decis Making 7, 351–360 (2008)
8. Huang, Z.X.: Some Properties of the Subspace on an L-fuzzy Order-preserving Operator Space. In: 2005 International General Topology Symposium, pp. 19–20. Zhangzhou Normal University, Zhangzhou (2005)
9. Chen, S.L., Cheng, J.S.: A New Hausdorff's Separation in Lω-spaces. Journal of Yangtze University 2(1), 8–13 (2005)
10. Chen, S.L.: ω-base and its Properties in L-fuzzy Order-preserving Operator Spaces. Journal of Jianghan Petroleum Institute 25(3), 143–145 (2003)

Using Two-Level Fuzzy Pattern Recognition in the Classification of Convex Quadrilateral

Xiang-jun Cui and Bing-yuan Cao*

School of Mathematics and Information Sciences,
Guangzhou University, Guangzhou 51006, China
e-mail: cuixiangjun159@163.com, bycao@gzhu.edu.cn

Abstract. This paper is devoted to the classification of convex quadrilateral by using fuzzy pattern recognition. Based on the principle of threshold and maximum membership, the two-level fuzzy pattern recognition is applied to solve this problem and a new model of membership function and the corresponding algorithms are given. We shall see that the obtained membership is advantageous for the recognition, and the final adjustment is reasonable. Moreover, the results obtained in this paper may be valuable and significant for the automatic recognition in some practical applications.

Keywords: Fuzzy pattern recognition, convex quadrilateral, membership function, threshold principle, maximum membership principle, algorithm.

1 Introduction

In the viewpoint of mathematics, convex quadrilateral contains trapezoidal, parallelogram and non-typical quadrangle in terms of the parallelism of the opposite sides in theory. The trapezoid divides into the isosceles trapezoid, the right-angle trapezoid and the atypical trapezoid according to its sizes of base angles. The parallelogram divides into rectangle, diamond, square and the atypical parallelogram according to its interior angle and the relation between the neighboring sides. Convex quadrilateral is widely used in various designs, such as circuit diagram, mechanical drawing, building plans, geographic information, process diagram, etc.. The automatic recognition is valuable for the application in areas. For example, the trapezoid is usual in the realistic craft. The cross section of the dovetail slot in machining is isosceles trapezoid. For factors in process, the two base angles are impossible to be just right, which involves the precise of the craft and the measuring system. They are thought to be quality in permissible variation, and unqualified if surpasses this scope. When we make the computer recognize the diagram of the across section of the dovetail slot the result must be “the wrong diagram”. However such trapezoidal chart is practical. It is realistic and valuable that how to distinguish the isosceles trapezoid and the right angle trapezoid in production. The similar parallelogram recognition examples can be seen everywhere.

* Corresponding author.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 527–534.
springerlink.com

It is necessary to regard them as fuzzy conception when carries on the pattern recognition. In 1976, it has been suggested to identify graphics by the fuzzy way [1-13], mainly on the triangular fuzzy pattern recognition (FPR). However, to the authors' knowledge, there are few relevant academic results for the identification of convex quadrilateral.

The main purpose of this paper is to improve the existing fuzzy graph identification model for the classification of convex quadrilateral. The rest of this paper is organized as follows: in Section 2, we will give some auxiliary definitions and basic formulas. In Section 3, we firstly describe how to conduct the two-level FPR, and then we shall establish a new model of membership function for FPR. In Section 4, we put forward the algorithms, and in Section 5, we give some examples to illustrate the applications of our abstract results. Finally, some concluding remarks are discussed.

2 Some Auxiliary Definitions and Basic Formulas

Definition 2.1. *(Maximum Membership Principle). Let U be the domain of discourse. Denote A as one of the fuzzy pattern included in U.* $u_1, u_2, \cdots, u_n$ *are belonged to U and* $u_1, u_2, \cdots, u_n$ *are the objects to be identified. If* $A(u_i) = \max\{A(u_1), A(u_2), \cdots, A(u_n)\}$, *then* u_i *is recognized to be belonged to* A .

Definition 2.2. *(Threshold Principle). Let U be the domain of discourse. Denote* $A_1, A_2, \cdots, A_P$ *as the fuzzy pattern included in U. Denote* λ *as a membership,* λ *is belonged to* $(0,1]$. u *is belonged to* U *and* u *is one of the objects to be identified. If* $\max\{A_1(u), A_2(u), \cdots, A_p(u)\} < \lambda$, *we refuse recognition, and find the reason; if* $\max\{A_1(u), A_2(u), \cdots, A_p(u)\} \geq \lambda$, *and there are k entries of fuzzy pattern* $A_{i1}(u)$, $A_{i2}(u), \cdots, A_{ik}(u)$, *which are greater than or equal to* λ, *identifying that u is belonged to* $\bigcap_{j=1}^{k} A_{ij}$.

Definition 2.3. *(The Two-level FPR). Let* $u_1, u_2, \cdots, u_n$ *be the objects to be identified.* $u_1, u_2, \cdots, u_n$ *are belonged to U. Denote* $A_i (i = 1,2,\cdots, p)$ *as the p entries of genre included in U.* A_i *is one of the genre, which includes q entries of classification marked to* $A_{ij} (i = 1,2,\cdots, p, j = 1,2,\cdots q)$. *Using the threshold principle and maximum membership principle, identifying that* u_i *belongs to* A_i , *then repeating again in the genre of* A_i, *and recognizing* u_i *belongs to* A_{ij} .

Definition 2.1-2.3 are classical, which can be found in many books and papers (e.g.[2-4]).

3 The Two-Level FPR for Convex Quadrilateral

In order to use the method of two-level FPR to solve the recognition classification of convex quadrilateral, we divide the discussion into five steps in the following.

3.1 Electing the Characteristics of the Objects

In the all of factors about u, elect the significant factors of the objects, which will be identified. Measuring their specific data, and then writing the index vector of the object recognition performance, we mark $u=(u_1,u_2,\cdots,u_m)$. For the Convex quadrilateral ABCD, it is determined by the four interior angles and the adjacent edge of the equivalence. Also the adjacent edge of the equivalence can be determined by recognizing the triangle ABC for its isosceles, so we elect the characteristics by the four interior angles, the angle BAC and angle BCA, which is marked by $u=(A,B,C,D,\angle BAC,\angle BCA)$.

3.2 Constructing the Membership Function of the First-Level FPR [5-6]

Constructing membership function is the keystone and difficult points. Because fuzzy pattern is the fuzzy sets of discourse domain, then constructing the membership function of fuzzy pattern is constructing membership function of fuzzy sets.

Let the four interior angles of convex quadrilateral to be marked by angle A, angle B, angle C and angle D. They determine the characters of convex quadrilateral, which contains the trapezoid, the parallelogram and the atypical quadrilateral. So let the discourse domain to be all of the convex quadrilateral, namely $U=\{(A,B,C,D)|A+B+C+D=360°,0°<A,B,C,D<180°\}$. Using the fuzzy sets A_T, A_P and A_{AQ} to denote the trapezoid, the parallelogram and the atypical quadrilateral. Since the genre of trapezoid, parallelogram and atypical quadrilateral is determined by the parallel of the two groups of subtense, which is equal to the complementary of the adjacent angles, and notice that if one group of subtense of convex quadrilateral is parallel and another is not parallel, it is belonged to trapezoid; if two groups of subtense are both parallel, it is belonged to parallelogram, the membership functions of fuzzy sets A_T,A_P,A_{AQ} are constructed in the following.

$$A_T=\begin{cases}1-\dfrac{\min\{|A+B-180°|,|B+C-180°|\}}{90°}, & others,\\ 0, & A+B=B+C=180°,\end{cases} \tag{1}$$

$$A_P=1-\frac{\max\{\ |A+B-180°|,|B+C-180°|\ \}}{180°}, \tag{2}$$

$$A_{AQ}=(1-A_T)\wedge(1-A_P). \tag{3}$$

3.3 Judgments of the First-Level FPR [7-9]

To solve the first-level FPR for convex quadrilateral, the corresponding membership function of the three genre should be computed. Using the given threshold λ_1 and maximum membership principle to recognize, chalk up the result, which is one of the trapezoid, parallelogram or atypical quadrilateral. If it is belonged to the trapezoid or parallelogram, the two-level FPR will be applied; if it is belonged to the atypical quadrilateral, FPR will be stopped, and the result will be marked by atypical quadrilateral.

3.4 Constructing the Membership Function of the Two-Level FPR

3.4.1 The Genre of Trapezoid [10-11]

When the convex quadrilateral is recognized to trapezoid under the first-level FPR, recognizing the isosceles, right-angled and atypical of the trapezoid is indispensable. Using the fuzzy sets A_{IT}, A_{RT} and A_{AT} to denote the isosceles trapezoid, right-angled trapezoid and atypical trapezoid. Because the class of isosceles, right-angled and atypical of the trapezoid is determined by the addend or minus of the degree and the right-angled of the base angle, their membership functions are constructed in the following.

$$A_{IT}=\begin{cases}1-\dfrac{\min\{\ |A-B|,|C-D|\ \}}{180°}, & |A+B-180°|>|B+C-180°|,\\[2ex] 1-\dfrac{\min\{\ |A-D|,|B-C|\ \}}{180°}, & |A+B-180°|\le|B+C-180°|,\end{cases} \tag{4}$$

$$A_{RT}=1-\frac{\min\{\ |A-90°|,|B-90°|,|C-90°|,|D-90°|\ \}}{90°}, \tag{5}$$

$$A_{AT}=(1-A_{IT})\wedge(1-A_{RT})\,. \tag{6}$$

3.4.2 The Genre of Parallelogram

When the convex quadrilateral is recognized to parallelogram under the first-level FPR, recognizing the rectangle, diamond, square and atypical of the parallelogram is determined by recognizing the triangle ABC for its isosceles. Using the fuzzy sets $A_{RP}, A_{DP}, A_{SP}, A_{AP}$ to denote the rectangle, diamond, square and atypical Parallelogram. The four membership functions are constructed in the following.

$$A_{RP} = 1 - \frac{\max\{\ |A-90°|, |B-90°|, |C-90°|, |D-90°|\ \}}{90°}, \tag{7}$$

$$A_{DP} = 1 - \frac{|\angle BAC - \angle BCA|}{60°}, \tag{8}$$

$$A_{SP} = A_{RP} \wedge A_{DP}, \tag{9}$$

$$A_{AP} = (1 - A_{RP}) \wedge (1 - A_{DP}) \wedge (1 - A_{SP}). \tag{10}$$

3.5 Judgments of the Two-Level FPR

To solve the two-level FPR for convex quadrilateral, when the convex quadrilateral is recognized to trapezoid under the first-level, for recognizing the isosceles, right-angled and atypical of the trapezoid, the corresponding membership function of the three class should be computed. Using the given threshold λ_2 and maximum membership principle to recognize, chalk up the result.

When the convex quadrilateral is recognized to parallelogram under the first-level, the corresponding membership function of the four classes of parallelogram should be computed. Using the given threshold λ_2 and maximum membership principle to recognize, chalk up the result.

4 The Algorithms

The four interior angles of the convex quadrilateral are angle A, angle B, angle C and angle D. They determine the characters of convex quadrilateral, which contains the trapezoid, the parallelogram and the atypical quadrilateral. When the first-level FPR identifying the convex quadrilateral is a parallelogram, based on recognizing its four interior angles and the angle BAC, angle BCA, according to recognize the triangle ABC, we can completely determine the classification of the convex quadrilateral. The basic algorithm steps of the two-level FPR for convex quadrilateral are summarized in the following.

Algorithm statement

Step 1: For a given convex quadrilateral ABCD, recognizing its four interior angles and the angle BAC, angle BCA;

Step 2: Enter into the first-level FPR, compute A_T, A_P and A_{AQ};

Step 3: Compare the $\max\{A_T, A_P\}$ and λ_1, if $\max\{A_T, A_P\} \geq \lambda_1$, then go to the next step; if $\max\{A_T, A_P\} < \lambda_1$, then go to Step 18;

Step 4: Compare the A_T and A_P, if $A_T > A_P$, it comes to the next step; if $A_T \leq A_P$, then go to Step 8;

Step 5: Enter into the two-level FPR, compute A_{IT}, A_{RT} and A_{AT} ;

Step 6: Compare the $\max\{A_{IT}, A_{RT}\}$ and λ_2, if $\max\{A_{IT}, A_{RT}\} \geq \lambda_2$, then go on; if $\max\{A_{IT}, A_{RT}\} < \lambda_2$, then go to Step 13;

Step 7: Compare the A_{IT} and A_{RT}, if $A_{IT} > A_{RT}$, then go to the Step 11; if $A_{IT} \leq A_{RT}$, then go to Step 12;

Step 8: Enter into the two-level FPR, compute the A_{RP}, A_{DP}, A_{SP} and A_{AP};

Step 9: Compare $\max\{A_{RP}, A_{DP}, A_{SP}\}$ and λ_2, if $\max\{A_{RP}, A_{DP}, A_{SP}\} \geq \lambda_2$, then it continues; if $\max\{A_{RP}, A_{DP}, A_{SP}\} < \lambda_2$, then go to Step 17;

Step 10: Compare the A_{RP}, A_{DP} and A_{SP}, if $A_{RP} > A_{DP} = A_{SP}$, then go to the step 14; if $A_{DP} > A_{RP} = A_{SP}$, then go to Step 15; if $A_{DP} > A_{RP} = A_{SP}$, then go to Step 16;

Step 11: Stop, output: convex quadrilateral ABCD is belonged to Isosceles Trapezoid.

Step 12: Stop, output: convex quadrilateral ABCD is belonged to Right-angled Trapezoid.

Step 13: Stop, output: convex quadrilateral ABCD is belonged to Atypical Trapezoid.

Step 14: Stop, output: convex quadrilateral ABCD is belonged to Rectangle.

Step 15: Stop, output: convex quadrilateral ABCD is belonged to Diamond.

Step 16: Stop, output: convex quadrilateral ABCD is belonged to Square.

Step 17: Stop, output: the quadrilateral ABCD is belonged to Atypical Parallelogram.

Step 18: Stop, output: the quadrilateral ABCD is belonged to Atypical Quadrilateral.

5 The Analysis of Numerical Experiments and Results

There are ten numerical examples. Their index vectors are given in the following.

$u_1 = (85°, 140°, 37°, 98°, 30°, 10°)$; $u_2 = (125°, 60°, 116°, 59°, 63°, 57°)$;
$u_3 = (15°, 70°, 160°, 115°, 8°, 102°)$; $u_4 = (138°, 140°, 37°, 45°, 25°, 15°)$;
$u_5 = (65°, 142°, 36°, 117°, 20°, 18°)$; $u_6 = (86°, 91°, 94°, 89°, 33°, 56°)$;
$u_7 = (87°, 90°, 96°, 87°, 47°, 43°)$; $u_8 = (124°, 60°, 117°, 59°, 80°, 40°)$;
$u_9 = (86°, 141°, 36°, 97°, 25°, 14°)$; $u_{10} = (110°, 75°, 101°, 74°, 53°, 52°)$.

Let λ_1 be equal to 0.8, and let λ_2 be equal to 0.9.

The following table summarizes the computational results on the above ten examples. In the table A_T, A_P, A_{AQ}, A_{IT}, A_{RT}, A_{AT}, A_{RP}, A_{DP}, A_{SP} and A_{AP} denote the degree of membership functions. The results show that our algorithm can globally solve the two-level FPR for the convex quadrilateral effectively.

Table 1. Comparison of recognizing results

No.	λ_1	λ_2	A_T	A_P	A_{AQ}	A_{IT}	A_{RT}	A_{AT}	A_{RP}	A_{DP}	A_{SP}	A_{AP}	Result
1	0.8	0.9	0.9667	0.7500	0.0333	0.6944	0.9444	0.0556					RT
2	0.8	0.9	0.9556	0.9722	0.0278				0.6111	0.9000	0.6111	0.1000	DP
3	0.8	0.9	0.4444	0.4722	0.5278								AQ
4	0.8	0.9	0.9667	0.4556	0.0333	0.9889	0.5000	0.0111					IT
5	0.8	0.9	0.9778	0.8500	0.0222	0.5722	0.7222	0.2778					AT
6	0.8	0.9	0.9667	0.9722	0.0278				0.9556	0.6167	0.6167	0.0444	RP
7	0.8	0.9	0.9667	0.9667	0.0333				0.9333	0.9333	0.9333	0.0667	SP
8	0.8	0.9	0.9667	0.9778	0.0222				0.6222	0.3333	0.3333	0.3778	AP
9	0.8	0.9	0.9667	0.7389	0.0333	0.6944	0.9556	0.0444					RT
10	0.8	0.9	0.9556	0.9722	0.0278				0.7778	0.9833	0.7778	0.0167	DP

6 Conclusion

This paper puts forward a more effective model of membership function, and gives the algorithms for the model of two-level FPR, analyzes the rationality and superior of the two-level FPR in the practical application of circle diagram, craft chart and so on. However, it is needed further research for progressive, advantage of the membership function and the abstained data is fuzzy convex interval, trapezoid fuzzy number [12-13], as well as polygons and other irregular shapes recognition.

Acknowledgements

This research was supported by the National Natural Science Foundation of China under Grant No. 70771030 and the National Natural Science Foundation of China under Grant No. 70271047.

References

1. Lee, E.T.: An application of fuzzy sets to the classification of geometric figures and chromosome images. Information Sciences 10, 95–114 (1976)
2. Cao, B.-y.: Applied fuzzy mathematics and systems. Science Press of China, Beijing (2005)
3. Cao, B.-y.: Pattern classification model with T-fuzzy data. Advances in Soft Computing 42, 793–802 (2007)

4. Chen, S.-l., Li, J.-g., Wang, X.-g.: Theory of fuzzy sets and its applications. Science Press of China, Beijing (2006)
5. Zhang, X.: Fuzzy pattern recognition of irregular triangles and square. Mechanical Science and Technology 10, 1145–1147 (2008)
6. Cui, X.-j.: A research of selecting the membership functions in the fuzzy pattern recognition. Journal of Shaoyang University 7(2), 5–8 (2010)
7. Wu, S.-l.: Popular fuzzy mathematics and programming. Water Resources and Hydropower Press of China, Beijing (2008)
8. Liang, B.-s., Cao, D.-l.: Fuzzy mathematics and its applications. Science Press of China, Beijing (2007)
9. Peng, Z.-z., Sun, Y.-y.: Applications of fuzzy mathematics. Wuhan University Press, Wuhan (2004)
10. Wei, Q.: Trapezoid categorization by pattern recognition. Journal of Qinghai Junior Teachers College 3, 18–19 (2000)
11. Cao, B.-y.: Fuzzy exponential model and its application in forecasting telephone amount. Journal of Shantou University(Natural Science Edition) 18(2), 1–6 (2003)
12. Cao, B.-y.: Interval and fuzzy convex function with convex functional. Journal of Guangzhou University (Natural Science Edition) 7(4), 1–5 (2008)
13. Cao, B.-y.: Cluster and recognition model with T-fuzzy data. Mathematical Theory and Applied Probability 11(4), 317–325 (1996)

Moore-Smith $\omega\theta$-Convergence Theory in ω-Molecular Lattices

Shui-li Chen[1,2], Jia-liang Xie[1,2], Guo-rong Cai[1,2], and Yun-dong Wu[1,2]

[1] School of Science, Jimei University, Xiamen,
Fujian 361021, P.R. China
[2] Research Centre of Image Information Engineering,
Jimei University, Xiamen, Fujian 361021, P.R. China
e-mail: sgzx@jmu.edu.cn, xiejialiang@jmu.edu.cn
e-mail: kiser@jmu.edu.cn, wuyundong@yahoo.com.cn

Abstract. In this paper, Moore-Smith $\omega\theta$-convergence theory of molecular nets and ideals in an ω-molecular lattice are established. By means of the $\omega\theta$-convergence theory, some important characterizations with respective to the $\omega\theta$-closed element and weakly (ω_1, ω_2)-continuous generalized order-homomorphisms are obtained.

Keywords: Fuzzy lattice, molecular net, ideal, generalized order-homomorphism, $\omega\theta$-convergence.

1 Introduction

The theory of topological molecular lattices, which is a generalization of fuzzy topology, was presented by Wang [8]. In order to unify various closure operators such as θ-closure operator [1], δ-closure operator [7], σ-closure operator [2] etc. in topological molecular lattices, a generalized molecular lattice which call an ω-molecular lattice was introduced by Chen [3]. Since then, a series of profound research works have been launched [4-6]. In this paper, we shall further enrich and consummate Moore-Smith convergence theory in ω-molecular lattices, and establish the Moore–Smith $\omega\theta$-convergence theory in ω-molecular lattices.

2 Preliminaries

Throughout the paper, L, L_1 and L_2 denote fuzzy lattices, while M, M_1 and M_2 denote the sets consisting of all molecules, i. e., nonzero $\vee$-irreducible elements in L, L_1 and L_2 respectively. 0 and 1 are the least and the greatest element of L respectively.

Definition 2.1. [9] *Let L be a complete lattice, $e \in L$, $B \subset L$. B is called a minimal family of e if $B \neq \emptyset$ and*

(i) $\sup B = e$;

(ii) $\forall A \subset L$, $\sup A \geq e$ implies that $\forall x \in B$, there exists $y \in A$ such that $y \geq x$.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing, AISC 82, pp. 535–543.
springerlink.com

According to Hutton [8], in a completely distributive lattice, each element $e\in L$ has a greatest minimal family which will be denoted by $\beta(e)$. For each $e\in M$, $\beta^*(e)=\beta(e)\cap M$ is a minimal family of e and is said to be the standard minimal family of e.

Definition 2.2. [3] *Let L be a fuzzy lattices.*

(i) An operator ω: $L\to L$ is said to be an ω-operator if (1) $\forall A, B\in L$ and $A\leq B$, $\omega(A)\leq\omega(B)$; (2) $\forall A\in L$, $A\leq\omega(A)$.

(ii) An element $A\in L$ is called an ω-element if $\omega(A)=A$.

(iii) Put $\Omega=\{A\in L|\ \omega(A)=A\}$, and call the pair (L,Ω) an ω-molecular lattice (briefly ω-ML).

Definition 2.3. [3] *Let (L,Ω) be an ω-ML, $P\in L$ and $e\in M$. If there exists $Q\in\Omega$ such that $e\not\leq Q$ and $P\leq Q$, then call P an ωR-neighborhood of e. The collection of all ωR-neighborhoods of e is denoted by $\omega\eta(e)$.*

Definition 2.4. [3] *Let (L,Ω) be an ω-ML, $A\in L$ and $e\in M$. If $A\not\leq P$ for each $P\in\omega\eta(e)$. Then e is said to be an ω-adherence point of A, and the union of all ω-adherence points of A is said to be the ω-closure of A, and is denoted by $\omega cl(A)$. If $A=\omega cl(A)$, then call A an ω-closed element in L. If A is an ω-closed element, then we say that A' is an ω-open element. If $P=\omega cl(P)$ and $e\not\leq P$, then P is said to be an ω-closed R-neighborhood (briefly, ωCR-neighborhood of e), and the collection of all ωCR-neighborhoods of e is denoted by $\omega\eta^{-}(e)$.*

Definition 2.5. [3] *Let (L,Ω) be an ω-ML, $A\in L$ and $\omega int(A)=\vee\{B\in L/B\leq A$ and B is an ω-open element$\}$. We call $\omega int(A)$ the ω-interior of A. Obviously, A is ω-open if and only if $A=\omega int(A)$.*

Definition 2.6. [5] *Let (L,Ω) be an ω-ML, $\mathcal{N}$ a molecular net in L and $e\in M$. Then:*

(i) e is said to be an ω-limit point of $\mathcal{N}$, or $\mathcal{N}$ ω-converges to e, in symbols, $\mathcal{N}\to_{\omega}e$, if $\mathcal{N}$ is eventually not in P for each $P\in\omega\eta^{-}(e)$. The union of all ω-limit points of $\mathcal{N}$ will be denoted by ω-lim$\mathcal{N}$.

(ii) e is said to be an ω-cluster point of $\mathcal{N}$, or $\mathcal{N}$ ω-accumulates to e, in symbols, $\mathcal{N}\propto_{\omega}e$, if $\mathcal{N}$ is frequently not in P for each $P\in\omega\eta^{-}(e)$. The union of all ω-cluster points of $\mathcal{N}$ will be denoted by ω-ad$\mathcal{N}$.

Proposition 2.1 [9] *Let L be a completely distributive lattice. Then each element of L is a union of some $\vee$-irreducible elements.*

3 $\omega\theta$-Convergence of Molecular Nets on ω-MLs

In this section, we shall present some concepts of $\omega\theta$-convergence of molecular nets in an ω-ML, and discuss their properties.

Definition 3.1. *Let (L,Ω) be an ω-ML, $\mathcal{N}$ a molecular net in L and $e\in M$. Then:*

(ii) e is said to be an $\omega\theta$-limit point of $\mathcal{N}$, or $\mathcal{N}$ $\omega\theta$-converges to e, in symbols, $\mathcal{N}\rightarrow_{\omega\theta}e$, if $\mathcal{N}$ is eventually not in $\omega int(P)$ for each $P\in\omega\eta^{-}(e)$. The union of all $\omega\theta$-limit points of $\mathcal{N}$ will be denoted by $\omega\theta$-lim$\mathcal{N}$.

(ii) e is said to be an $\omega\theta$-cluster point of $\mathcal{N}$, or $\mathcal{N}$ $\omega\theta$-accumulates to e, in symbols, $\mathcal{N}\propto_{\omega\theta}e$, if $\mathcal{N}$ is frequently not in $\omega int(P)$ for each $P\in\omega\eta^{-}(e)$. The union of all $\omega\theta$-cluster points of $\mathcal{N}$ will be denoted by $\omega\theta$-ad$\mathcal{N}$.

Theorem 3.1. *Let (L,Ω) be an ω-ML, $e\in M$ and let $\mathcal{N}$ be a molecular net in L. Then:*
(1) $\mathcal{N}\rightarrow_{\omega\theta}e$ if and only if $\mathcal{N}\rightarrow_{\omega\theta}b$ for each $b\in\beta^*(e)$;
(2) $\mathcal{N}\propto_{\omega\theta}e$ if and only if $\mathcal{S}\propto_{\omega\theta}b$ for each $b\in\beta^*(e)$.

Proof. (1) Suppose that $\mathcal{N}\rightarrow_{\omega\theta}e$ and $b\in\beta^*(e)$. Then $b\leq e$. Hence $P\in\omega\eta^{-}(e)$ for $P\in\omega\eta^{-}(b)$, and hence $\mathcal{N}$ is eventually not in $\omega int(P)$ for each $P\in\omega\eta^{-}(b)$ by $\mathcal{N}\rightarrow_{\omega\theta}e$. Conversely, assume that $\mathcal{N}\rightarrow_{\omega\theta}b$ for each $b\in\beta^*(e)$. If e is not an $\omega\theta$-limit point of $\mathcal{N}$, then there exists $P\in\omega\eta^{-}(e)$ such that $\mathcal{N}$ is frequently in $\omega int(P)$. Since $e=\sup\beta^*(e)$, there is $d\in\beta^*(e)$ with $d\nleq P$, that is, $P\in\omega\eta^{-}(d)$. This means that d is not an $\omega\theta$-limit point of $\mathcal{N}$. Hence, the sufficiency is proved.

(2) Similar to the proof of (1).

Theorem 3.2. *Let (L,Ω) be an ω-ML, $e\in M$ and let $\mathcal{N}$ be a molecular net in L. Then*:

(1) $\mathcal{N}\rightarrow_{\omega\theta}e$ if and only if $e\leq\omega\theta$-lim$\mathcal{N}$;
(2) $\mathcal{N}\propto_{\omega\theta}e$ if and only if $e\leq\omega\theta$-ad$\mathcal{N}$;
(3) $\omega\theta$-lim$\mathcal{N}\leq\omega\theta$-ad$\mathcal{N}$.

Proof. We only check (1), the proofs of (2) and (3) are omitted. If $\mathcal{N}\rightarrow_{\omega\theta}e$, then $e\leq\omega\theta$-lim$\mathcal{N}$ by the definition of $\omega\theta$-lim$\mathcal{N}$. Conversely, if $e\leq\omega\theta$-lim$\mathcal{N}$, then for each $b\in\beta^*(e)$, there exists an $\omega\theta$-limit point d of $\mathcal{N}$ with $b\leq d$ by virtue of the fact that $e=\sup\beta^*(e)$ and the definition of $\omega\theta$-lim$\mathcal{N}$. Since $P\in\omega\eta^{-}(d)$ for $P\in\omega\eta^{-}(b)$, $\mathcal{N}$ is eventually not in $\omega int(P)$ by $\mathcal{N}\rightarrow_{\omega\theta}d$. Consequently, $\mathcal{N}\rightarrow_{\omega\theta}e$ according to Theorem 3.1.

Definition 3.2. *Let (L,Ω) be an ω-ML, $A\in L$ and $e\in M$. If $A\nleq\omega int(P)$ for each $P\in\omega\eta(e)$, then e is said to be an $\omega\theta$-adherence point of A, and the union of all $\omega\theta$-adherence points of A is said to be the $\omega\theta$-closure of A, and is denoted by $\omega\theta cl(A)$. If $A=\omega\theta cl(A)$, then call A an $\omega\theta$-closed element in L. If A is an $\omega\theta$-closed element, then we say that A' is an $\omega\theta$-open element.*

Theorem 3.3. *Let* (L,Ω) *be an* ω*-ML,* $A\in L$ *and* $e\in M$*. Then e is an* $\omega\theta$*-adherence point of A if and only if there is a molecular net* $\mathfrak{N}$ *in A such that* $\mathfrak{N}\to_{\omega\theta}e$*.*

Proof. If e is an $\omega\theta$-adherence point of A, then according to Definition 3.2, $A\not\leq \omega int(P)$ for each $P\in \omega\eta^{-}(e)$. Taking $\mathfrak{N}=\{N(P)\in A \mid A\not\leq \omega int(P),\ P\in \omega\eta^{-}(e)\}$, then $\mathfrak{N}$ is a molecular net in A because $\omega\eta^{-}(e)$ is a directed set and $\mathfrak{N}\to_{\omega\theta}e$. Conversely, if there is a molecular net $\mathfrak{N}=\{N(n)\mid n\in D\}$ in A such that $\mathfrak{N}\to_{\omega\theta}e$, then $\mathfrak{N}$ is eventually not in $\omega int(P)$ for each $P\in \omega\eta^{-}(e)$. Because $N(n)\in A$ for each $n\in D$, $A\not\leq \omega int(P)$ for each $P\in \omega\eta^{-}(e)$. This implies that e is an $\omega\theta$-adherence point of A.

Theorem 3.4. *Let* $\mathfrak{N}=\{N(n)\mid n\in D\}$ *be a molecular net in* (L,Ω)*,* $e\in M$ *and* $\mathfrak{N}\to_{\omega\theta}e$*. Then* $\mathfrak{T}\to_{\omega\theta}e$ *for each subnet* $\mathfrak{T}$ *of* $\mathfrak{N}$*.*

Proof. Assume that $\mathfrak{T}=\{T(m)\mid m\in E\}$ is a subnet of $\mathfrak{N}$. By the definition of subnet, there exists a mapping R: $E\to D$ satisfying the following conditions: (1) $\forall\ m\in E$, $T(m)= N(R(m))$; (2) $\forall\ n_0\in D$, there exists $m_0\in E$ such that $R(m)\geq n_0$ whenever $m\geq m_0$. Since $\mathfrak{N}\to_{\omega\theta}e$, for each $P\in \omega\eta^{-}(e)$ we can choose $n_0\in D$ such that $N(n)\not\leq \omega int(P)$ whenever $n\geq n_0$. According to (2), there exists $m_0\in E$ satisfying $R(m)\geq n_0$ as $m\geq m_0$. Hence $T(m)= N(R(m))\not\leq \omega int(P)$ for each $P\in \omega\eta^{-}(e)$. This shows that $\mathfrak{T}\to_{\omega\theta}e$.

Theorem 3.5. *Let* $\mathfrak{N}$ *be a molecular net in* (L,Ω) *and* $e\in M$*. Then e is an* $\omega\theta$*-cluster point of* $\mathfrak{N}$ *if and only if there exists a subnet* $\mathfrak{T}$ *of* $\mathfrak{N}$ *which e is an* $\omega\theta$*-limit point of* $\mathfrak{T}$*.*

Proof. Assume that e is an $\omega\theta$-cluster point of $\mathfrak{N}=\{N(n)\mid n\in D\}$. By Definition 3.1(ii), $\mathfrak{N}$ is frequently not in $\omega int(P)$ for each $P\in \omega\eta^{-}(e)$, i.e., there exists $m\in E$ such that $m\geq n$ and $N(m)\not\leq \omega int(P)$ for each $n\in D$. Write $m=R(n,P)$ and $E=\{R(n,P)\mid (n,P)\in D\times \omega\eta^{-}(e)\}$. Define a binary relation "$\leq$" in E as follows: $\forall\ R(n_1,P_1), R(n_2,P_2)\in E$,

$$R(n_1,P_1)\leq R(n_2,P_2) \text{ if and only if } n_1\leq n_2 \text{ and } P_1\leq P_2.$$

Then E is a directed set about the relation "$\leq$". Taking $T(m)=N(R(n,P))$, one can easily see that $\mathfrak{T}=\{T(m)\mid m\in E\}$ is a subnet of $\mathfrak{N}$ and e is an $\omega\theta$-limit point of $\mathfrak{T}$.

Conversely, suppose that $\mathfrak{T}=\{T(m)\mid m\in E\}$ is a subnet of $\mathfrak{N}$ and that e is an $\omega\theta$-limit point of $\mathfrak{T}$. According to the definition of subnet, there exists a mapping R: $E\to D$ and $m_0\in E$ such that $R(m)\geq n_0$ whenever $m\geq m_0$ for each $n_0\in D$. Since e is an $\omega\theta$-limit point of $\mathfrak{T}$, there is $m_1\in E$ such that $T(m)\not\leq \omega int(P)$ as $m\geq m_1$ for each $P\in \omega\eta^{-}(e)$. Since E is a directed set, we can choose $m_2\in E$ with $m_2\geq m_0$ and $m_2\geq m_1$, thus $T(m_2) \nleq \omega int(P)$ and $R(m_2) \geq n_0$. Taking $n=R(m_2)$, then $n\in D$ and $N(n)=N(R(m_2)) \nleq \omega int(P)$ for each $P\in \omega\eta^{-}(e)$. Hence, e is an $\omega\theta$-cluster point of $\mathfrak{N}$.

Theorem 3.6. *Let (L,Ω) be an ω-ML and $A\in L$. Then A is an $\omega\theta$-closed element if and only if for each molecular net $\mathfrak{N}$ in A, $\omega\theta$-lim$\mathfrak{N} \leq A$.*

Proof. Let A be an $\omega\theta$-closed element and $\mathfrak{N}$ a molecular net in A. If $e\leq\omega\theta$-lim$\mathfrak{N}$, then e is an $\omega\theta$-adherence point of A according to Theorem 3.3, that is, $e\leq\omega\theta\mathrm{cl}(A)=A$. Therefore, $\omega\theta$-lim$\mathfrak{N}\leq A$ by the arbitrariness of e in $\omega\theta$-lim$\mathfrak{N}$. Conversely, if $e\leq\omega\theta\mathrm{cl}(A)$, then there exists a molecular net $\mathfrak{N}$ in A satisfying $e\leq\omega\theta$-lim$\mathfrak{N}$ by Theorem 3.3 and Theorem 3.2. Hence $e\leq\omega\theta$-lim$\mathfrak{N} \leq A$ in line with the sufficient condition. This means that $\omega\theta\mathrm{cl}(A)\leq A$, i.e., A is $\omega\theta$-closed.

Theorem 3.7. *If $\mathfrak{N}$ be a molecular net in (L,Ω). Then* $\omega\theta$-lim$\mathfrak{N}$ *and* $\omega\theta$-ad$\mathfrak{N}$ *are both ω-closed elements in (L,Ω).*

Proof. For each $e\in M$, if $e\leq\omega\mathrm{cl}(\omega$-lim$\mathfrak{N})$, then $\omega\theta$-lim$\mathfrak{N} \nleq P$ for each $P\in \omega\eta^{-}(e)$ by Definition 2.4. With reference to Proposition 2.1 we can choose a molecule $b\leq\omega\theta$-lim$\mathfrak{N}$ with $b\nleq P$, i.e., $P\in \omega\eta^{-}(b)$. Consequently, $\mathfrak{N}$ is eventually not in $\omega int(P)$ for each $P\in \omega\eta^{-}(e)$ by $b\leq\omega\theta$-lim$\mathfrak{N}$. This shows that $e\leq\omega\theta$-lim$\mathfrak{N}$, and thus $\omega\mathrm{cl}(\omega\theta$-lim$\mathfrak{N})\leq\omega\theta$-lim$\mathfrak{N}$ by the arbitrariness of e in $\omega\mathrm{cl}(\omega\theta$-lim$\mathfrak{N})$. On the other hand, $\omega\theta$-lim$\mathfrak{N}\leq\omega\mathrm{cl}(\omega\theta$-lim$\mathfrak{N})$ by Theorem 2.5 in [3]. Therefore, $\omega\theta$-lim$\mathfrak{N}$ is ω-closed.

Similarly, for each $e\leq\omega\mathrm{cl}(\omega\theta$-ad$\mathfrak{N})$ we have $\omega\theta$-ad$\mathfrak{N} \nleq P$ for each $P\in \omega\eta^{-}(e)$ by Definition 2.4, that is, there is a molecule $d\leq\omega\theta$-ad$\mathfrak{N}$ with $d\nleq P$, i.e., $P\in \omega\eta^{-}(d)$. Hence, $\mathfrak{N}$ is frequently not in $\omega int(P)$ for each $P\in \omega\eta^{-}(e)$ in accordance with $d\leq\omega\theta$-ad$\mathfrak{N}$, and thus $e\leq\omega\theta$-ad$\mathfrak{N}$. So, $\omega\mathrm{cl}(\omega\theta$-ad$\mathfrak{N})\leq\omega\theta$-ad$\mathfrak{N}$ by the arbitrariness of e in $\omega\mathrm{cl}(\omega\theta$-ad$\mathfrak{N})$. It implies that $\omega\theta$-ad$\mathfrak{N}$ is also an ω-closed element in (L,Ω).

4 $\omega\theta$-Convergence of Ideals on ω-MLs

In this section, we shall present some concepts of $\omega\theta$-convergence of ideals in an ω-ML, and discuss their properties.

Definition 4.1. *Let (L,Ω) be an ω-ML, $\mathscr{I}$ an ideal in L and $e\in M$. Then:*

(i) e is said to be an $\omega\theta$-limit point of $\mathscr{I}$, or $\mathscr{I}$ $\omega\theta$-converges to e, in symbols, $\mathscr{I}\to_{\omega\theta}e$, if $\omega int(P)\in\mathscr{I}$ for each $P\in \omega\eta^{-}(e)$. The union of all $\omega\theta$-limit points of $\mathscr{I}$ will be denoted by $\omega\theta$-lim$\mathscr{I}$.

(ii) e is said to be an $\omega\theta$-cluster point of $\mathscr{I}$, or $\mathscr{I}$ $\omega\theta$-accumulates to e, in symbols, $\mathscr{I}\propto_{\omega\theta}e$, if $B\vee\omega int(P)\neq 1$ for each $P\in\omega\eta^{-}(e)$ and each $B\in\mathscr{I}$. The union of all $\omega\theta$-cluster points of $\mathscr{I}$ will be denoted by $\omega\theta$-ad$\mathscr{I}$.

Similar to the $\omega\theta$-convergence of molecular nets, we can obtain the following results for $\omega\theta$-convergence of ideals.

Theorem 4.1. *Let (L,Ω) be an ω-ML, $e\in M$ and let $\mathscr{I}$ be an ideal in L. Then*:

(1) $\mathscr{I}\to_{\omega\theta}e$ if and only if $\mathscr{I}\to_{\omega\theta}b$ for each $b\in\beta^*(e)$;
(2) $\mathscr{I}\propto_{\omega\theta}e$ if and only if $\mathscr{I}\propto_{\omega\theta}b$ for each $b\in\beta^*(e)$.

Theorem 4.2. *Let (L,Ω) be an ω-ML, $e\in M$ and let $\mathscr{I}$ be an ideal in L. Then*:

(1) $\mathscr{I}\to_{\omega\theta}e$ if and only if $e\leq\omega\theta$-lim$\mathscr{I}$;
(2) $\mathscr{I}\propto_{\omega\theta}e$ if and only if $e\leq\omega\theta$-ad$\mathscr{I}$;
(3) $\omega\theta$-lim$\mathscr{I}\leq\omega\theta$-ad$\mathscr{I}$.

Theorem 4.3. *Suppose that $\mathscr{I}_1$ and $\mathscr{I}_2$ are two ideals in an ω-ML (L,Ω) which $\mathscr{I}_1\subset\mathscr{I}_2$. Then*:

(1) if e is an $\omega\theta$-limit point of $\mathscr{I}_1$, then e is also an $\omega\theta$-limit point of $\mathscr{I}_2$;
(2) if e is an $\omega\theta$-cluster point of $\mathscr{I}_2$, then e is also an $\omega\theta$-cluster point of $\mathscr{I}_1$.

Theorem 4.4. *Let (L,Ω) be an ω-ML, $A\in L$ and $e\in M$. Then e is an $\omega\theta$-adherence point of A if and only if there is an ideal $\mathscr{I}$ in L which $A\notin\mathscr{I}$ and $\mathscr{I}\to_{\omega\theta}e$.*

Theorem 4.5. *Let $\mathscr{I}$ be an ideal in an ω-ML (L,Ω) and $e\in M$. Then e is an $\omega\theta$-cluster point of $\mathscr{I}$ if and only if there exists an ideal $\mathscr{I}_1$ in L such that $\mathscr{I}\subset\mathscr{I}_1$ and e is an $\omega\theta$-limit point of $\mathscr{I}_1$.*

Theorem 4.6. *Let (L,Ω) be an ω-ML and $A\in L$. Then A is an $\omega\theta$-closed element if and only if for each ideal $\mathscr{I}$ in L such that $A\notin\mathscr{I}$ and $\omega\theta$-lim$\mathscr{I}\leq A$.*

Theorem 4.7. *If $\mathscr{I}$ be an ideal in an ω-ML (L,Ω), then $\omega\theta$-lim$\mathscr{I}$ and $\omega\theta$-ad$\mathscr{I}$ are both ω-closed elements in (L,Ω).*

Theorem 4.8. *Let $\mathscr{I}$ be a maximal ideal in an ω-ML (L,Ω). Then $\omega\theta$-lim$\mathscr{I}=\omega\theta$-ad$\mathscr{I}$.*

5 The Relationships between $\omega\theta$-Convergence of Molecular Nets and Ideals

In this section, we shall discuss the relationships between $\omega\theta$-convergence of molecular nets and ideals in ω-MLs.

Definition 5.1. [10] *Let $\mathcal{I}$ be an ideal in an ω-ML (L,Ω) and $D(\mathcal{I})=\{(b,B)\,/\,b\in M, B\in\mathcal{I}$ and $b\not\leq B\}$. In $D(\mathcal{I})$, define a binary relation "$\leq$" as follows: $\forall (b_1,B_1)$, $(b_2, B_2)\in D(\mathcal{I})$, $(b_1, B_1)\leq(b_2, B_2)$ if and only if $B_1\leq B_2$. Obviously $D(\mathcal{I})$ is a directed set equipped with the relation. Take $\mathcal{N}(\mathcal{I})=\{\mathcal{N}(\mathcal{I})\,(b, B)=b\,/(b, B)\in D(\mathcal{I})\}$. Then $\mathcal{N}(\mathcal{I})$ is a molecular net in L, we call $\mathcal{N}(\mathcal{I})$ the molecular net induced by $\mathcal{I}$.*

Theorem 5.1. *Let $\mathcal{I}$ be an ideal in an ω-ML (L,Ω). Then:*

(1) $\mathcal{I}\to_{\omega\theta}e$ if and only if $\mathcal{N}(\mathcal{I})\to_{\omega\theta}e$;

(2) $\mathcal{I}\propto_{\omega\theta}e$ if and only if $\mathcal{N}(\mathcal{I})\propto_{\omega\theta}e$.

Proof. (1) If $\mathcal{I}\to_{\omega\theta}e$, then $\omega int(P)\in\mathcal{I}$ for each $P\in\omega\eta^{-}(e)$ and hence $(e, P)\in D(\mathcal{I})$. According to Definition 5.1, $P\leq B$ for each $(b, B)\in D(\mathcal{I})$ and $(e, P)\leq(b, B)$. Therefore, $\mathcal{N}(\mathcal{I})\,(b, B)\not\leq \omega int(P)$ by $\mathcal{N}(\mathcal{I})(b, B)\not\leq B$. This means that e is an $\omega\theta$-limit point of $\mathcal{N}(\mathcal{I})$. Conversely, if $\mathcal{N}(\mathcal{I})\to_{\omega\theta}e$, then there exists an element $(a, A)\in D(\mathcal{I})$ such that $\mathcal{N}(\mathcal{I})(b, B)\not\leq\omega int(P)$ whenever $(a, A)\leq(b, B)$ for each $P\in\omega\eta^{-}(e)$ in accordance with Definition 3.1. Specially, choose $A=B$, we have $\mathcal{N}(\mathcal{I})(b, A)\not\leq\omega int(P)$ from $\mathcal{N}(\mathcal{I})(b, A)\not\leq A$, i.e., $\omega int(P)\leq A$. Since $\mathcal{I}$ is a lower set, $\omega int(P)\in\mathcal{I}$ by $A\in\mathcal{I}$. Consequently, e is an $\omega\theta$-limit point of $\mathcal{I}$.

(2) Assume $\mathcal{I}\propto_{\omega\theta}e$, in line with Definition 4.1, $B\vee\omega int(P)\neq 1$ for each $P\in\omega\eta^{-}(e)$ and each $B\in\mathcal{I}$. From Proposition 2.1 we can take a molecule $b\in M$ with $b\not\leq B\vee\omega int(P)$, thus $(b, B)\in D(\mathcal{I})$ and $\mathcal{N}(\mathcal{I})(b, B)\not\leq\omega int(P)$. It shows that $\mathcal{N}(\mathcal{I})$ is frequently not in $\omega int(P)$ for each $P\in\omega\eta^{-}(e)$, i.e., $\mathcal{N}(\mathcal{I})\propto_{\omega\theta}e$. Conversely, if $\mathcal{N}(\mathcal{I})\propto_{\omega\theta}e$, then there is an element $(a, A)\in D(\mathcal{I})$ with $(b, B)\leq(a, A)$ such that $\mathcal{N}(\mathcal{I})(a, A)\not\leq\omega int(P)$ for each $P\in\omega\eta^{-}(e)$ and each $(b, B)\in D(\mathcal{I})$. Since $B\leq A$ and $a\not\leq A$, we know that $a\not\leq B$, and so $B\vee\omega int(P)\neq 1$. It is implies that e is an $\omega\theta$-cluster point of $\mathcal{I}$, that is, $\mathcal{I}\propto_{\omega\theta}e$.

Definition 5.2. [10] *Let $\mathcal{N}$ be a molecular net in (L,Ω) and $\mathcal{I}(\mathcal{N})=\{B\in L\,|\,\mathcal{N}$ is eventually not in $B\}$. Then $\mathcal{I}(\mathcal{N})$ is an ideal in L, we call $\mathcal{I}(\mathcal{N})$ the ideal induced by $\mathcal{N}$.*

Theorem 5.2. *Let $\mathcal{N}$ be a molecular net in an ω-ML (L,Ω). Then*:

(1) $\mathcal{N}\to_{\omega\theta}e$ if and only if $\mathcal{I}(\mathcal{N})\to_{\omega\theta}e$;

(2) if $\mathcal{N}\propto_{\omega\theta}e$, then $\mathcal{I}(\mathcal{N})\propto_{\omega\theta}e$.

Proof. (1) If $\mathfrak{N}\to_{\omega\theta}e$, then $\mathfrak{N}$ is eventually not in $\omega int(P)$ for each $P\in \omega\eta^{-}(e)$, and then $\omega int(P)\in \mathfrak{I}(\mathfrak{N})$ by Definition 5.2. Hence, $\mathfrak{I}(\mathfrak{N})\to_{\omega\theta}e$. Conversely, assume that $\mathfrak{I}(\mathfrak{N})\to_{\omega\theta}e$. Then $\omega int(P)\in \mathfrak{I}(\mathfrak{N})$ for each $P\in \omega\eta^{-}(e)$, i.e., $\mathfrak{N}$ is eventually not in $\omega int(P)$ for each $P\in \omega\eta^{-}(e)$. Therefore $\mathfrak{N}\to_{\omega\theta}e$.

(2) If $\mathfrak{N}\propto_{\omega\theta}e$, then $\mathfrak{N}$ is frequently not in $\omega int(P)$ for each $P\in \omega\eta^{-}(e)$. According to Definition 5.2, $\mathfrak{N}$ is eventually not in B for each $B\in \mathfrak{I}(\mathfrak{N})$. This implies that $\mathfrak{N}$ is frequently not in $B\vee \omega int(P)$ for each $B\in \mathfrak{I}(\mathfrak{N})$ and each $P\in \omega\eta^{-}(e)$, that is, $B\vee \omega int(P)\neq 1$. Hence $\mathfrak{I}(\mathfrak{N})\propto_{\omega\theta}e$ from Definition 4.1.

6 Some Applications of $\omega\theta$-Convergence Theory in ω-MLs

In this section, we shall give some characterizations of weak (ω_1,ω_2)-continuity of generalized order-homomorphisms by means of $\omega\theta$-convergence theory of molecular nets and ideals.

Definition 6.1. *Let (L_i,Ω_i) be an ω_i-ML (i=1, 2) and $f:L_1\to L_2$ a generalized order-homomorphism (briefly, GOH[8]).*

(i) f is called weakly (ω_1, ω_2)-continuous if $f(\omega_1 cl(B))\leq \omega_2\theta cl f(A)$ for each $A\in L_1$.
(ii) f is called weakly (ω_1, ω_2)-continuous at $e\in M_1$ if $\omega_1 cl(f^{-1}(\omega_2 int(Q)))\in \omega_1\eta^{-}(e)$ for each $Q\in \omega_2\eta^{-}(f(e))$.

Obviously, f is weakly (ω_1,ω_2)-continuous if and only if for each $e\in M_1$, f is weakly (ω_1,ω_2)-continuous at e.

Theorem 6.1. *Let f be a GOH from an ω_1-ML (L_1,Ω_1) into an ω_2-ML (L_2,Ω_2). Then f is weakly (ω_1,ω_2)-continuous if and only if for every molecular net $\mathfrak{N}$ in L_1, $f(\omega_1\text{-}\lim\mathfrak{N})\leq \omega_2\theta\text{-}\lim f(\mathfrak{N})$.*

Proof. Assume that f is weakly (ω_1,ω_2)-continuous, $\mathfrak{N}=\{N(n)\mid n\in D\}$ is a molecular net in L_1, and $d\leq f(\omega_1\text{-}\lim\mathfrak{N})$. Then $f(\mathfrak{N})=\{f(N(n))\mid n\in D\}$ is a molecular net in L_2 and $\omega_1\mathrm{cl}(f^{-1}(\omega_2\mathrm{int}(Q)))\in \omega_1\eta^{-}(e)$ for each $Q\in \omega_2\eta^{-}(f(e))$. In accordance with the definition of GOH, there is a molecule $e\in M_1$ satisfying $d=f(e)$ and $e\leq\omega_1\text{-}\lim\mathfrak{N}$. Hence there exists $n_0\in D$ such that $N(n)\not\leq \omega_1\mathrm{cl}(f^{-1}(\omega_2\mathrm{int}(Q)))$ whenever $n\geq n_0$. Since $N(n)\not\leq \omega_1\mathrm{cl}(f^{-1}(\omega_2\mathrm{int}(Q)))$ implies $N(n)\not\leq f^{-1}(\omega_2\mathrm{int}(Q))$, $f(N(n))\not\leq \omega_2\mathrm{int}(Q)$. This shows $f(e)$ is an ω_2-limit point of $f(\mathfrak{N})$, i.e., $d\leq \omega_2\theta\text{-}\lim f(\mathfrak{N})$. Therefore, $f(\omega_1\text{-}\lim\mathfrak{N})\leq \omega_2\theta\text{-}\lim f(\mathfrak{N})$ by randomicity of d in $f(\omega_1\text{-}\lim\mathfrak{N})$.

Conversely, if f is not weakly (ω_1,ω_2)-continuous, then there is an element A in L_1 such that $f(\omega_1\mathrm{cl}(A))\not\leq \omega_2\theta\mathrm{cl}(f(A))$. From Proposition 2.1, we can choose $e\in M_1$, $e\leq \omega_1\mathrm{cl}(A)$ and $f(e)\leq f(\omega_1\mathrm{cl}(A))$, but $f(e)\not\leq \omega_2\theta\mathrm{cl}(f(A))$. There exists a molecular net $\mathfrak{N}$ in A such that e is an ω_1-limit point of $\mathfrak{N}$ from Theorem 3.4. Hence $f(e)$ is an ω_2-limit point of $f(\mathfrak{N})$, where $f(\mathfrak{N})=\{f(N(n))\mid n\in D\}$ is a molecular net in $f(A)$, and hence $f(e)\leq\omega_2\theta\text{-}\lim f(\mathfrak{N})$ by the hypothesis of sufficiency. Since $\omega_2\theta\mathrm{cl}(f(A))$ is

ω_2-closed, $f(e) \le \omega_2$-lim $f(\mathcal{N}) \le \omega_2\theta\mathrm{cl}(f(A))$ by Theorem 3.6. It is a contradiction with $f(e) \not\le \omega_2\theta\mathrm{cl}(f(A))$. Therefore f is weakly (ω_1, ω_2)-continuous.

Theorem 6.2. *Let f be a GOH from an ω_1-ML (L_1, Ω_1) into an ω_2-ML (L_2, Ω_2). Then f is weakly (ω_1, ω_2)-continuous if and only if for every ideal $\mathcal{I}$ in L_1, $f(\omega_1\text{-lim}\mathcal{I}) \le \omega_2\text{-lim}(f(\mathcal{I}'))'$.*

Proof. It follows from Definition 4.1 and Theorem 4.4.

7 Conclusion

In this paper, we establish the Moore–Smith $\omega\theta$-convergence theory of molecular nets and ideals in ω-MLs. As an application, we obtain some characterizations of weak (ω_1, ω_2)-continuity of generalized order-homomorphisms by means of the $\omega\theta$-convergence theory of molecular nets and ideals in ω-MLs.

Acknowledgements. The paper is supported by the Natural Science Foundation of Fujian Province (A0810014), Special fund for colleges and universities in Fujian Province Science (JK2009017) and Technology Agency, and Xiamen Municipal Science and Technology Plan Project (3502Z20093018) of China.

References

[1] Chen, S.L.: θ-convergence theory of ideals on topological molecular lattices. J. Math. 13(2), 163–169 (1993)
[2] Chen, S.L., Chen, S.-T., Wang, X.-G.: σ-convergence theory and its applications in fuzzy lattices. Information Sciences 165(1), 45–58 (2004)
[3] Chen, S.L.: On ω-molecular lattices. In: Proc. ICMLC 2005, vol. I, pp. 2619–2623 (2005)
[4] Chen, S.L.: Separation axioms in ω-molecular lattices. In: Proc. ICMLC 2006, vol. I, pp. 1801–1804 (2006)
[5] Chen, S.L.: ω-convergence theory of molecular nets in ω-molecular lattices. In: Proc. ICNC-FSKD 2010, vol. I (in press, 2010)
[6] Chen, S.L., Wu, Y.D., et al.: ω-Convergence theory of ideals in ω-molecular lattices. In: Proc. ICMLC 2010, vol. I (in press, 2010)
[7] Cheng, J.S.: Some properties of δ-continuous order-homomorphisms. Fuzzy Systems and Math. 11(4), 38–41 (1997)
[8] Hutton, B.: Uniformities on fuzzy topological spaces. J. Math. Anal. Appl. 58, 559–571 (1977)
[9] Wang, G.J.: Theory of topological molecular lattices. Fuzzy Sets and Systems 47, 351–376 (1992)
[10] Yang, Z.Q.: Ideals in topological molecular lattices. Acta Math. Sinica 29(2), 276–279 (1988)

Some Properties about $L\omega$-Quotient Spaces and $L\omega$-Product Spaces

Yu-qun Zeng

School of Sciences, Jimei University, Xiamen, 361021, China
e-mail: yqzeng@jmu.edu.cn

Abstract. In this paper, relations between order-preserving operators and closure operators in $L\omega$-spaces are explored. A necessary and sufficient condition about closure operators is obtained. Some properties about $L\omega$-quotient spaces and $L\omega$-product spaces are given.

Keywords: $L\omega$-spaces, ω-set, ω-closed set, order-preserving operator, $L\omega$-quotient space, $L\omega$-product space.

1 Introduction

The concept of L-fuzzy order-preserving operator spaces (briefly, $L\omega$-spaces) is presented in [2]. In this paper, we explore relations between order-preserving operator and closure operator in $L\omega$- spaces, and obtain some properties about $L\omega$-spaces, quotient spaces and $L\omega$-product spaces.

Definition 1.1. [2] *Let L be a fuzzy lattice, X be a non-empty crisp set. An L-fuzzy operator $\omega: L^X \to L^X$ is order-preserving , briefly said to be an Lω-operator if*

(1) $\forall A, B \in L^X$ *and* $A \le B$, $\omega(A) \le \omega(B)$.

(2) $\forall A \in L^X$, $A \le \omega(A)$.

Put $\Omega = \{A \in L^X \mid A = \omega(A)\}$, we call the pair (L^X, Ω) an $L\omega$-space. If $Q \in \Omega$, then call Q an ω-set in (L^X, Ω).

Definition 1.2. [2] *Let* (L^X, Ω) *be an Lω-space,* $P \in L^X$ *and* $x_\alpha \in M^*(L^X)$. *If there exists* $Q \in \Omega$ *such that* $P \le Q$ *and* $x_\alpha \not\le Q$, *then call* P *an* ωR*-neighborhood of* x_α. *The collection of all* ωR*-neighborhoods of* x_α *is denoted by* $\omega\eta(x_\alpha)$.

Definition 1.3. [2] *Let* (L^X, Ω) *be an Lω-space,* $A \in L^X$ *and* $x_\alpha \in M^*(L^X)$. *If* $\forall P \in \omega\eta(x_\alpha)$, *we have* $A \not\le P$, *then we call* x_α *an* ω*-adherence point of* A. *The union of all* ω*-adherence points of* A *is said to be the* ω*-closure of* A, *and is*

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 545–552.
springerlink.com

denoted by ω-cl(A). If A=ω-cl(A), then we call A *an* ω*-closed set in* (L^X,Ω)*, and call* A *an* ω*-open set in* (L^X,Ω)*. The operator* ω*-cl:* $L^X \to L^X$ *is said to be the* ω*-closure operator in* (L^X,Ω)*. If* P *is an* ω*--closed set and* $x_\alpha \nleqslant P$*, then call* P *an* ω*-closed R-neighborhood of* x_α*. The collection of all* ω*-closed R-neighborhoods of* x_α *is denoted by* $\omega\eta^-(x_\alpha)$*.*

2 Some Properties of ω-Sets

The following theorem is obvious.

Theorem 2.1. *Let* (L^X,Ω) *be an Lω-space. The closure operator in* (L^X,Ω) *is an order-preserving operator.*

Theorem 2.2. *Let* (L^X,Ω) *be an Lω-space and* $A \in L^X$*. Then* $\omega(A) \leqslant \omega$*-cl(A).*

Proof. Grant that x_α is a molecule and $x_\alpha \le \omega(A)$. $\forall P \in \omega\eta(x_\alpha)$, there exists $Q \in \Omega$ such that $P \le Q$ and $x_\alpha \nleq Q$, thus $\omega(A) \nleq Q$. We say $A \nleq P$. Otherwise, $A \le P$, we have $\omega(A) \le \omega(P)$. Since $\omega(P) \le \omega(Q) = Q$, we have $\omega(A) \le Q$. It is in contradiction with $\omega(A) \nleq Q$. Consequently $A \nleq P$. This shows that x_α is an ω-adherence point of A. Therefore $\omega(A) \leqslant \omega$-$cl(A)$.

Theorem 2.3. *Let* (L^X,Ω) *be an Lω-space and* $A \in L^X$*. Then* A *is an* ω*-set if and only if* A *is an* ω*-closed set.*

Proof. Assume that A is an ω-closed set. By Definition 1.3, $A = \omega$-$cl(A)$. According to Theorem2.2, we have $A \le \omega(A) \le \omega$-$cl(A)$. Hence $A = \omega(A)$, i.e., $A \in \Omega$.

Conversely, assume that A is an ω-set and that x_α be an ω-adherence point of A, i.e., $x_\alpha \le \omega$-$cl(A)$. We assert that $x_\alpha \le A$. In fact, if $x_\alpha \nleq A$, then A is an ω-R-neighborhood of x_α since $A \in \Omega$. Therefore, $\forall P \in \omega\eta(x_\alpha)$, $A \nleq P$. Specially, $A \nleq A$. This is impossible. Consequently, each ω-adherence point of A must be in A. In other words, if ω-$cl(A) \le A$, then ω-$cl(A)$=A, and then A is an ω-closed set according to definition 1.3.

It is easy to verify the following theorem.

Theorem 2.4. 0_X *is an* ω*-set in* (L^X,Ω)*.*

Theorem 2.5. *Let* (L^X,Ω) *be an Lω-space and* T *be an index set. If* $a_i \in \Omega$ *for each* $t \in T$*, then* $\wedge_{t\in T} A_t \in \Omega$.

Proof. Let $\wedge_{t\in T} A_t = B$. Suppose $x_\alpha \le \omega\text{-}cl(B)$. In other words, x_α is an ω-adherence point of *B*. Then we have $B \not\le P$ for $\forall P \in \omega\eta(x_\alpha)$. Hence $A_t \not\le P$ for each $t \in T$. This means that x_α is an ω-adherence point of A_t. Therefore, $x_\alpha \le \omega\text{-}cl(A_t)$. We can get $A_t = \omega\text{-}cl(A_t)$ with the hypothesis of $A_t \in \Omega$. Hence $x_\alpha \le A_t$ for each $t \in T$, i.e., $x_\alpha \le \wedge_{t\in T} A_t = B$. Thus $\omega\text{-}cl(B) \le B$, it implies that *B* is an ω-closed set. Consequently, $\wedge_{t\in T} A_t \in \Omega$ by theorem 2.3.

Theorem 2.6. *Let* (L^X, Ω) *be an Lω-space. If* $A \in \Omega$ *and* $B \in \Omega$, *then* $A \vee B \in \Omega$.

Proof. If $A \in \Omega$ *and* $B \in \Omega$, *then* $A = \omega\text{-}cl(A)$ *and* $B = \omega\text{-}cl(B)$ *by Theorem 2.3. We have* $A \vee B = \omega\text{-}cl(A) \vee \omega\text{-}cl(B) = \omega\text{-}cl(A \vee B)$. *Hence,* $A \vee B$ *is an* ω*-closed set, and hence* $A \vee B \in \Omega$ *by Theorem 2.3.*

Theorem 2.7. $\omega\eta^-(x_\alpha)$ *is an ideal base in* L^X.

Proof. Obviously, $1_X \in \omega\eta^-(x_\alpha)$. Assume that $Q_1 \in \omega\eta^-(x_\alpha)$ and $Q_2 \in \omega\eta^-(x_\alpha)$. Then $x_\alpha \not\le Q_1$ and $x_\alpha \not\le Q_2$. We assert that $x_\alpha \not\le Q_1 \vee Q_2$ because x_α is a molecule. Hence $Q_1 \vee Q_2 \in \omega\eta^-(x_\alpha)$. This shows that $\omega\eta^-(x_\alpha)$ is a directed set. Therefore $\omega\eta^-(x_\alpha)$ is an ideal base in L^X.

Similar to preceding theorem, it is easy to verify the following result.

Theorem 2.8. $\omega\eta(x_\alpha)$ is an ideal in L^X.

Theorem 2.9. *Let* (L^X, Ω) *be an Lω-space, and* $\Omega' = \{A' \in L^X \mid A \in \Omega\}$. *Then* (L^X, Ω') *is an L-topological space.*

Proof. we need to prove the following statements.

(1) $1_x \in \Omega'$, $0_X \in \Omega'$. We can easily verify that $1_X \in \Omega'$ by Theorem 2.4. According to Definition 1.1 and $1_X \in \Omega$, $0_X \in \Omega'$.

(2) Intersection of finite sets in Ω' is in Ω'. We can easily get the result from Theorem 2.6.

(3) Union of arbitrary sets in Ω' is in Ω'. The statement holds in accordance with theorem 2.5.

Therefore (L^X, Ω') is an *L*-topological space on L^X. So (L^X, Ω) is an L-cotopological space on X.

Theorem 2.10. *Let* (L^X, Ω) *be an Lω-space and* $A \in L^X$. *Then* ω*-cl(A)=*$\omega(A)$ *if and only if* $\omega(A)$ *is an* ω*-set.*

Proof. Assume that $\omega\text{-}cl(A)=\omega(A)$. Then $\omega\text{-}cl(\omega\text{-}cl(A))=\omega\text{-}cl(\omega(A))$. Since $\omega\text{-}cl(\omega\text{-}cl(A))=\omega\text{-}cl(A)$, we can get $\omega\text{-}cl(\omega(A))=\omega\text{-}cl(A)=\omega(A)$. This shows that $\omega(A)$ is an ω-closed set. Therefore, $\omega(A)$ is an ω-set. Conversely, assume that $\omega(A)$ is an ω-set, x_α is a molecule and $x_\alpha \leq \omega\text{-}cl(A)$. Then for $\forall P \in \omega\eta(x_\alpha)$, we have $A \nleq P$. We assert that $x_\alpha \leq \omega(A)$. Otherwise, $x_\alpha \nleq \omega(A)$. Since $\omega(A)$ is an ω-set, we can get $\omega(A) \in \omega\eta(x_\alpha)$. Hence , $A \nleq \omega(A)$ which contradicts Definition 1.1. Therefore, $\omega\text{-}cl(A) \leq \omega(A)$. Thus $\omega\text{-}cl(A)=\omega(A)$.

Theorem 2.11. *Let* (L^X,Ω) *be an* $L\omega$*-space,* $A \in L^X$ *and* $B \in L^X$. *Then*

$$\omega\text{-}cl(A \vee B) = \omega\text{-}cl(A) \vee \omega\text{-}cl(B).$$

Proof. We can get $\omega\text{-}cl(A) \vee \omega\text{-}cl(B) \leq \omega\text{-}cl(A \vee B)$ from Theorem2.1. On the other hand, for each molecule $x_\alpha \leq \omega\text{-}cl(A \vee B)$, we assert that $x_\alpha \leq \omega\text{-}cl(A) \vee \omega\text{-}cl(B)$. Otherwise, Suppose that $x_\alpha \nleq \omega\text{-}cl(A) \vee \omega\text{-}cl(B)$. Then $x_\alpha \nleq \omega\text{-}cl(A)$ and $x_\alpha \nleq \omega\text{-}cl(B)$. Hence there exist $P_1 \in \omega\eta(x_\alpha)$ and $P_2 \in \omega\eta(x_\alpha)$ satisfying $A \leq P_1$ and $B \leq P_2$. Obviously, $A \vee B \leq P_1 \vee P_2$ and $P_1 \vee P_2 \in \omega\eta(x_\alpha)$. Thus x_α is not an ω-adherence point of $A \vee B$. It is in contradiction with the hypothesis that $x_\alpha \leq \omega\text{-}cl(A \vee B)$. Hence, for each molecule $x_\alpha \leq \omega\text{-}cl(A \vee B)$, we have $x_\alpha \leq \omega\text{-}cl(A) \vee \omega\text{-}cl(B)$, i.e. $\omega\text{-}cl(A \vee B) \leq \omega\text{-}cl(A) \vee \omega\text{-}cl(B)$. Consequently, $\omega\text{-}cl(A \vee B) = \omega\text{-}cl(A) \vee \omega\text{-}cl(B)$.

Theorem 2.12. *Let* (L^X,Ω) *be an* $L\omega$*-space and* $A \in L^X$. *Then*

$$\omega\text{-}cl(A) = \wedge\{B \in L^X \mid B \in \Omega, A \leq B\}.$$

Proof. Since $\omega\text{-}cl(\omega\text{-}cl(A)) = \omega\text{-}cl(A)$, we get that $\omega\text{-}cl(A)$ is an ω-closed set. In the light of Theorem2.3, $\omega\text{-}cl(A)$ is an ω-set. If B is an ω-set and $A \leq B$, then $\omega\text{-}cl(A) \leq B$. Thus $\omega\text{-}cl(A) \leq \wedge\{B \in L^X \mid B \in \Omega, A \leq B\}$. On the other hand, $A \leq \omega\text{-}cl(A)$, and $\omega\text{-}cl(A)$ is an ω-set. Hence $\omega\text{-}cl(A) = \wedge\{B \in L^X \mid B \in \Omega, A \leq B\}$.

Theorem 2.13. *Let L be a fuzzy lattice, X be a non-empty crisp set. An operator* $\omega : L^X \to L^X$, *if*

(1) $\forall A, B \in L^X$ *and* $A \leq B$, $\omega(A) \leq \omega(B)$;
(2) $\forall A \in L^X$, $A \leq \omega(A)$;
(3) $\omega(\omega(A))=\omega(A)$.

Then $\delta = \{A \in L^X \mid A' = \omega(A')\}$ *is an L-topology on X. Moreover,* $\forall A \in L^X$, $\omega(A) = \wedge\{B \in L^X \mid B' \in \delta, A \leq B\}$, *i.e.* ω *is a closure operator.*

Proof. $\omega(0_X)=0_X$ follows from Theorem 2.4, and δ is an L-topology on X by Theorem 2.9. We can verify $\omega(A\vee B)=\omega(A)\vee\omega(B)$ by Theorem 2.10 and 2.11. By Theorem 2.12 and by Theorem 2.1.10 in [1] under conditions (2) and (3), we can verify that theorem 2.13 holds.

3 Quotient Space and Its Properties

Definition 3.1. *Let (L_1^X,Ω_1) be an $L\omega$-space, L_2 an fuzzy lattice, Y a non-empty crisp set, $f:L_1^X \to L_2^Y$ is an order-epimorphism and $\Omega_2=\{B\in L_2^Y \mid f^{-1}(B)\in\Omega_1\}$. We call the $L_2\omega$-space (L_2^Y,Ω_2) is the quotient space of (L_1^X,Ω_1) about f. f is quotient order-homomorphism from (L_1^X,Ω_1) to (L_2^Y,Ω_2).*

Let $\Omega_2'=\{B'\in L_1^X \mid f^{-1}(B)\in\Omega_1\}$.We can easily verify Ω_2' is an L-topology on L_2^Y, and Ω_2' is the finest topology on which f is continuous.

Theorem 3.1. *Let (L_2^Y,Ω_2) be the quotient space of (L_1^X,Ω_1) about the order-epimorphism f, and (L_3^Z,Ω_3) be an $L_3\omega$-space, $g:L_2^Y\to L_3^Z$ be an order-homomorphism. g is continuous if and only if the composite order-homomorphism $g\circ f:L_1^X\to L_3^Z$ is continuous.*

Proof. Since f is quotient order-homomorphism and continuous order-homomorphism. Hence if $g:L_2^Y\to L_3^Z$ is continuous, then we can easily get the composite order-homomorphism $g\circ f:L_1^X\to L_3^Z$ is continuous.

On the other hand, if the composite order-homomorphism $g\circ f:L_1^X\to L_3^Z$ is continuous, then $(g\circ f)^{-1}(A)\in\Omega_1$ for $\forall A\in\Omega_3$, i.e. $f^{-1}(g^{-1}(A))\in\Omega_1$. We get $g^{-1}(A)\in\Omega_2$ by the definition of Ω_2. Consequently, $g:L_2^Y\to L_3^Z$ is continuous.

Theorem 3.2. *Let (L_1^X,Ω_1) and (L_2^y,Ω_2) be both order-preserving operator spaces, $f:L_1^X\to L_2^Y$ is continuous order-epimorphism. If f is closed (open) order-homomorphism, then f is quotient order-homomorphism.*

Proof. Let $\Omega_2=\{B\in L_1^X \mid f^{-1}(B)\in\Omega_1\}$. Assume that f is closed order-homomorphism. We only need to prove that $\Omega=\Omega_2$.

If $B\in\Omega_2$, then we have $f^{-1}(B)\in Q_1$ by the definition of Ω_2. Since f is closed order-homomorphism, so we have $ff^{-1}(B)\in\Omega$. Since f is full, we have $ff^{-1}(B)=B$. Thus $B\in\Omega$.

If $B\in\Omega$, then we have $f^{-1}(B)\in\Omega_1$ by continuity of f. We can get $B\in\Omega_2$ by the definition of Ω_2.

Consequently, $\Omega=\Omega_2$, and f is quotient order-homomorphism.

In the same way , we can get f is quotient order-homomorphism if f is open order-homomorphism.

Theorem 3.3. *Let* (L_1^X,Ω_1) *and* (L_2^Y,Ω_2) *be both order-preserving operator spaces,* $f: L_1^X \to L_2^Y$ *be a faithful and full and continuous order-homomorphism.* f *is closed if and only if* f *is the quotient order-homomorphism.*

Proof. Assume that f is closed order-homomorphism. We can easily get f is the quotient order-homomorphism from theorem3.2. Now, consider the case that f is the quotient order-homomorphism. For each $B \in \Omega_1$, since f faithful, we have $f^{-1}f(B) = B$, hence $f^{-1}f(B) \in \Omega_1$. We can get $f(B) \in \Omega_2$ by the definition of Ω_2. Consequently, f is closed.

4 Product Space and Its Properties

Definition 4.1. *Let* $\{(L^{X_t},\Omega_t)\}_{t\in T}$ *be a collection of* $L\omega$*-spaces,* T *a nonempty index set,* $X = \prod_{t\in T} X_t$ *the direct product of nonempty crisp sets* $\{X_t\}$*. The Zadeh's type function* $P_t : L^X \to L^{X_t}$ *is induced by the usual projective mapping* $P_t: X \to X_t$*. We also call the* $P_t : L^X \to L^{X_t}$ *projective mapping. Let* $\gamma = \{P_t^{-1}(A_t) \mid A_t' \in \Omega_t, t \in T\}$*, then* γ *is subbase of* Ω'*.* Ω' *generated by* γ *is L-topology on* L^X*. We call* (L^X,Ω) *the product space of* $\{(L^{X_t},\Omega_t)\}_{t\in T}$*.*

Definition 4.2. *Let* (L^X,Ω) *be an* ω*-space. If each LF-set* $[\lambda]$ *which is constant value* λ *for* $\forall \lambda \in L$ *is an* $L\omega$*-closed set, then we call* (L^X,Ω) *a stratified* ω*-space.*

Theorem 4.1. *Let* (L^X,Ω) *be the product space of* $\{(L^{X_t},\Omega_t)\}_{t\in T}$*. The projective m apping* $P_t : L^X \to L^{X_t}$ *is continuous order homomorphism for* $\forall t \in T$*.*

Proof. Since $P_t : L^X \to L^{X_t}$ is an Zadeh's type function induced by the projective mapping $P_t : X \to X_t$, hence it is order homomorphism. Since $P_t^{-1}(A_t) \in \gamma \subseteq \Omega'$ for arbitrary $A_t \in \Omega'_t$, we have $P_t^{-1}(A_t) \in \Omega'$, therefore the projective mapping $P_t : L^X \to L^{X_t}$ is continuous.

Theorem 4.2. *Let* (L^X,Ω) *be the product space of* $\{(L^{X_t},\Omega_t)\}_{t\in T}$*. If* (L^{X_τ},Ω_τ) *is a stratified* $L\omega$*-space for some* $\tau \in T$*, then the projective mapping* $P_\tau : L^X \to L^{X_\tau}$ *is open order homomorphism .*

Proof. Let $\beta = \{\wedge_{t\in S} P_t^{-1}(A_t) \mid A_t' \in \Omega_t, S \in 2^{(T)}\}$, where $2^{(T)}$ represents the set composed of all nonempty finite intersection of T, then β composes base of

Ω'.Let $A = \underset{t\in S}{\wedge}\{P_t^{-1}(A_t) \mid A_t' \in \Omega_t\}$ for arbitrary member A in base β, we assert that $P_t A$ is an ω-open set in L^{X_τ}, therefore the projective mapping $P_\tau : L^X \to L^{X_\tau}$ is open order homomorphism .

If $S = \{\tau\}$, then $A = P_\tau^{-1}(A_\tau) \in \gamma \subseteq \Omega'$ where $A'_\tau \in \Omega_\tau$. Since $P_t A = A$, therefore $P_t A$ is an ω-open set in L^{X_τ}.

Let $x \in X$, $P_\tau(x) = a$, $A(x) = \underset{t\in S}{\wedge} P_t^{-1}(A_t)(x) = \underset{t\in S}{\wedge}\{P_t^{-1}(A_t)(x)\} = \underset{t\in S}{\wedge} A_t(P_t(x))$. If $\tau \notin S$, then $A(x)$ is not relational to a. Since $P_\tau A(a) = \underset{x\in X}{\vee}\{A(x) \mid P_\tau(x) = a\}$, therefore $P_t A(a)$ is not relational to a. This shows that $P_\tau(A)$ is constant value. We can easily get that $P_\tau(A)$ is an ω-open set because $(L^{X_\tau}, \Omega_\tau)$ is a stratified $L\omega$-space.

If $\tau \in S$ and $S - \{\tau\}$ nonempty, we denote $B = \underset{t\in S-\{\tau\}}{\wedge}\{P_t^{-1}(A_t) \mid A_t' \in \Omega_t\}$, then $P_t B$ is an ω-open set in L^{X_τ}, and $A = P_\tau^{-1}(A_\tau) \wedge B$ is an ω-open set. $\forall a \in X_t$,

$$
\begin{aligned}
P_\tau A(a) &= P_\tau\{P_\tau^{-1}(A_\tau) \wedge B\}(a) \\
&= \vee\{\{P_\tau^{-1}(A_\tau) \wedge B\}(x) \mid P_\tau(x) = a,\ x \in X\}\ \vee\{P_\tau^{-1}(A_\tau)(x) \wedge B(x) \mid P_\tau(x) = a, x \in X\} \\
&= \vee\{A_\tau(a) \wedge B(x) \mid P_\tau(x) = a, x \in X\} \\
&= A_\tau(a) \wedge \{\vee B(x) \mid P_\tau(x) = a, x \in X\} \\
&= A_\tau(a) \wedge P_\tau B(a) = (A_\tau \wedge P_\tau B)(a),
\end{aligned}
$$

this shows that $P_t A = A_t \wedge P_t B$. Both A_t and $P_t B$ are ω-open sets in L^{X_τ}, therefore $P_t A$ is an ω-open set in L^{X_τ}.

Consequently, the projective mapping $P_\tau : L^X \to L^{X_\tau}$ is an open order homomorphism.

Lastly, it is easy to verify the following result.

Theorem 4.3. *Let (L^X, Ω) be the product space of $\{(L^{X_t}, \Omega_t)\}_{t\in T}$, $(L_1^{\ Y}, \mu)$ an L-topological space, and $f : L_1^{\ Y} \to L^X$ an order homomorphism. Then f is continuous if and only if $P_t \circ f$ is continuous for $\forall t \in T$.*

Acknowledgements. The paper is supported by the Natural Science Foundation of Fujian Province of China (A0810014).

References

1. Wang, G.J.: Thoery of L-fuzzy topological spaces. Shanxi Normal University Press (1988)
2. Chen, S.L.: On L-fuzzy Order-preserving Operator ω-spaces. The Journal of Fuzzy Mathematics 14(2), 481–499 (2006)
3. Chen, S.L.: Several order-homomorphisms on L–fuzzy topological spaces. Shanxi Normal University 16(3), 15–19 (1988)

4. Chen, S.L.: ω-base and Its Properties in L-fuzzy Order-preserving Operator Spaces. Journal of Jianghan Petroleum Institute 25(3), 143–145 (2003)
5. Chen, S.L., Wu, Y.D., Cai, G.R.: Omega Theta-Convegence Theory of Filters in L -Spaces. In: Proc. FSKD (IEEE), vol. 7, pp. 683–687 (2008)
6. Chen, S.L., Chen, G.L., Zhang, J.C.: ω-Convergence Theory of Filters in L -Spaces. Fuzzy Optimization and Decision Making 7(4), 351–360 (2008)
7. Chen, S.L., Wu, Y.D., Cai, G.R., Xie, J.L.: The almost ω-compactness in L ω-Spaces. Advances of Soft Computing 62, 1701–1710 (2009)
8. Chen, S.L.: Some separation axioms L ω-Spaces. In: Proc. AFSA World Congress, vol. I (2006)
9. Huang, Z.X.: The connectedness on L-fuzzy order-preserving operator spaces. In: Proceeding of International Conference on FuzzyInformation Processing Theories and University Press&Springer, pp. 165–168 (2003)

Self-adaptive Fuzzy PID Controller for Airborne Three-Axis Pan-Tilt

Yun-dong Wu[1,3], Li-chuan Geng[3], Qiang Zhang[2], and Shui-li Chen[3]

[1] Research Center of Image Information Engineering,
Jimei University, Amoy 361021, China
[2] Information Engineering University, Zhengzhou 450052, China
[3] College of Sciences, Jimei University, Amoy 361021, China
e-mail: wuyundong@yahoo.com.cn

Abstract. Airborne three-axis pan-tilt is one key equipment of Low-altitude Unmanned Aerial Vehicle Photogrammetry System (LUAVPS). In order to improve the control precision of the pan tilt, a self-adaptive fuzzy PID controller is proposed, combining traditional PID control and Fuzzy control technology. This PID controller can adaptively complete PID parameters adjustment, realizing pan-tilt self stabilization. Simulation results show that this fuzzy PID controller has high stable state precision and small hysteresis quality, pan-tilt control precision of LUAVPS can be fully fulfilled.

Keywords: Self-adaptive fuzzy PID, pan-tilt control, UAV, low-altitude photogrammetry.

1 Introduction

According to “Topographic Map Aerial Photogrammetry Standard”, the camera primary optical axis must keep vertical, and the row or column direction of image plane must point to north at the photo moment. All the exterior angle orientation parameters must almost equal to zero. But during the photo taking process, UAV’s pitch, yaw and roll angel are changing momentarily because of the change of wind direction, power and flight path. So a self-stabilization pan-tilt is needed between the UAV and camera to keep the exterior angle orientation parameters almost equal to zero.

Three-axis self stabilization pan-tilt is a steady output system, it’s a nonlinear system. PID control and Fuzzy control are all efficient tools to settle nonlinear system control. A good pan-tilt must have small steady state error and short delay time, but steady-state error and delay time are two contradictions in one system. When steady state is good, delay time maybe long. On the contrary, when delay time is short, the steady state maybe worse.

Conventional PID controller has better steady state, but long delay time, fuzzy controller has worse steady state but short delay time[1-2]. According to this, a self-adaptive fuzz PID controller is presented, which combining the advantage of

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 553–559.
springerlink.com

PID controller and Fuzzy controller, and the performance of the UAV's pan-tilt system is improved.

2 Design and Realization of Fuzzy PID Controller

2.1 Three Axis Self-Stabilization Pan-tilt

This pan-tilt is controlled by three channel controllers, each channel controller controls one axis to keep three exterior angel orientation parameters of photogrammetry coordinate system close to zero which caused by UAV's pitch, yaw and roll. The camera's state during flying is obtained by the gyroscope. The whole pan-tilt control process is to keep the image exterior orientation parameters close to zero by adjusting the three channel controllers. Figure.1 shows the control principle of UAV's pan-tilt system.

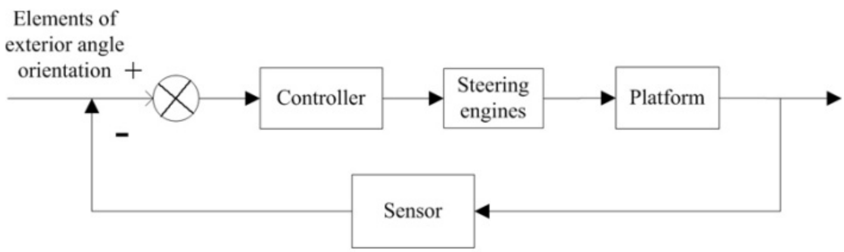


Fig. 1. Control principle of UAV's pan-tilt system

Because three channel controllers have the same control requirement, just one channel controller is discussed in this paper.

2.2 Traditional PID Control

The basic formula of PID controller is as follows [3-5]:

$$u(t) = K_P[e(t) + \frac{1}{T_1}\int_0^t e(t)dt + T_D\frac{de(t)}{dt}] \tag{1}$$

in which K_p is the proportional gain, T_1 is the integrate constant, T_D is the derivate constant, $u(t)$ is the analog control variable, $e(t)$ is the control error.

This control method is based on the precise model of the controlled object, the method is simple, practical and easy to realize, and it's always used in time invariant systems. There are several short comes is practice:

1) Practice system always complex, nonlinear, time variability and incompleteness, so it's hard to acquire the precise math model;

2) Pointing to the practical systems, harsh linear hypothesis is always needed, but these assumptions always don't match the actual system;
3) Actual control systems are always complex, and the traditional control task demand is low, facing complex control task, PID controller could do nothing.

Under the influence of noise and load perturbation, UAV pan-tilt system process parameter even model structure will change along time and working conditions. It's very difficult to find a group of appropriate PID parameter to suit the system's wide scope regulation. This requires the PID parameter adjusting process doesn't depend on the system's mathematic model, and the PID parameters can realize self adjustment.

2.3 Traditional Fuzzy Control

In practical project, a complex control system may obtain satisfy control effect by an experienced operator. This shows that if the human mind can be simulated by a controller, complex systems control can be realized, this results the fuzzy control.

Fuzzy control is one intelligent control method which based on fuzzy set theory, fuzzy language variable and fuzzy logic inference. We will not introduce Fuzzy logic in detail; it can be seen in [9]. It imitates human's behavior of fuzzy reasoning and decision making process. First, the operator or expert's experience is compiled into fuzzy rules, and then real-time signal from the sensor is changed to fuzzy signal, the fuzzy signal as the input of fuzzy rules, fuzzy reasoning is accomplished, the reasoning output is added in the operator. Figure 2 shows the basic theory.

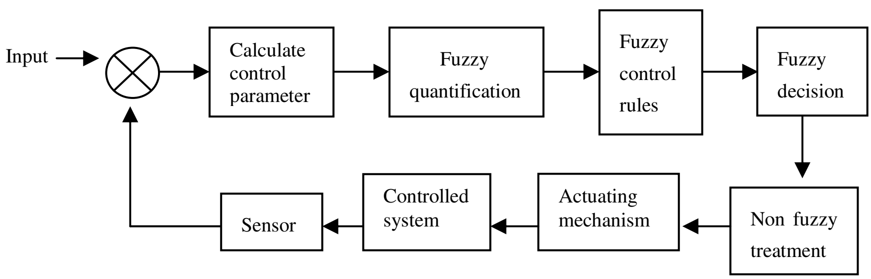


Fig. 2. Fuzzy control basic theory

Fuzzy control is suitable to industrial processes and large system control, the more difficult of the establishment of the control system's mathematical model, the better its superiority over other control methods will be reflected.

Fuzzy control theory has several characters:

1) Fuzzy control doesn't need the object's precise model;
2) Fuzzy control is an intelligent method reflecting human's wisdom which can be easily accepted;
3) Fuzzy control is robust and adaptive.

But comparing to the traditional control theory, fuzzy control is still developing, its theory and method are not perfect, even seems immature. Because it's a non-linear control method, rule explosion problem exists, so the control table or the control analytical formula can't be too large or complex. Fuzzy control system is actually a non-linear P or PD control method, static error exists in theory without imports in integrates mechanism. When the control rule structure and coverage is improper, or the selection of proportional factor or quantitative factor is improper, the system apt to produce vibrates, especially when the central language variable range is improper.

2.4 Structure and Adjusting Principle of Fuzzy PID Controller

PID controller and fuzzy controller both have its limitation in reality, several improved PID controllers are developed, but the complex control task remain can't be fully fulfilled. The inconvenience of fuzzy control method which caused by its nonlinear limits its application in various control tasks.

In this paper, pointing to the shortcomings of the traditional PID controller and fuzzy control, fuzzy control and conventional PID control are combined, A PID parameter self adjusting controller is designed, K_P , K_I and K_D are adjusted on real-time according to the change and change rate of the exterior orientation parameters. The input variation is the exterior angle orientation parameter φ and its change rate φ_c , the output is the three parameters of the PID controller. In order to fulfill the self adjusting requirement of φ and φ_c , the three PID controller parameters are changed online using the fuzzy control rules, and this composed the fuzzy PID controller. As shown in figure 3.

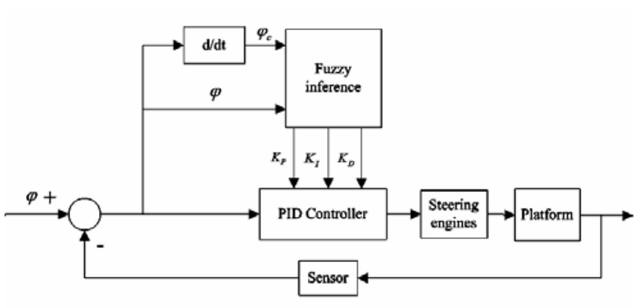


Fig. 3. Fuzzy PID controller

From figure 3 we can find that this controller is composed with traditional PID controller, fuzzy reasoning and its parameters. Fuzzy reasoning and parameter refinement is a fuzzy controller in reality. Its input is φ and φ_c and the output is the three PID parameters K_P , K_I and K_D .

The realization thought is to find the fuzzy relation between K_P, K_I, K_D and φ, φ_c .During flight φ and φ_c are detected continually to fulfill the requirement of φ and φ_c to the control parameters in different time.

Considering the pan-tilt's stability, response speed, stable accuracy and the PID parameters function between each other. To different $|\varphi|$ and $|\varphi_c|$, self adjusting requirement of the controlled process to K_P, K_I, K_D can be concluded as bellow [6-8]:

1) When $|\varphi|$ is large, K_P must be bigger so to speed up system response speed. At the same time in order to avoid control effect exceeding permitted extent, when $|\varphi|$ changed with time and differential oversaturation appears, K_D must be small. K_I should equal to zero to avoid large overcontrol of system response.

2) When $|\varphi|$ and $|\varphi_c|$ are middle, K_P should be a little smaller, so to acquire small over-control, K_I should be appropriate, the value of K_D has big influence to system response, its value should be middle so to guarantee system response speed.

3) When $|\varphi|$ is small or almost equal to zero, K_P and K_I should be increased to keep system stability. Considering system anti-interference performance, K_D is very important to avoid system oscillation near the set value. When $|\varphi_c|$ is small, K_D should be a little bigger. When $|\varphi_c|$ is big, K_D should be a little smaller. When $|\varphi_c|$ is big, K_P should be small, K_I should be big.

3 Establishment of Membership Function and Fuzzy Control Rules

The system input are φ and φ_c, their fuzzy language value has seven grades, denote as $\{NB, NM, NS, ZE, PS, PM, PB\}$ .The output fuzzy language value has four grades, denote as $\{Z, S, M, B\}$. All the input membership functions are gaussmf, outputs are trimf. Membership function of φ is shown in figure 4.

Discourse domain of φ is $[-180, +180]$, φ_c is $[-100, +100]$, K_P, K_I, K_D is $[0,150]$, $[0,5]$, $[0,15]$ .The fuzzy control rules are established according self adjusting rules in 2.4 and real pan-tilt control experience. As shown in table 1, The first rule is to say when φ is NB and φ_c is also NB , K_P is *B*, K_I is *Z*, K_D is *S*, other rules are all the same.

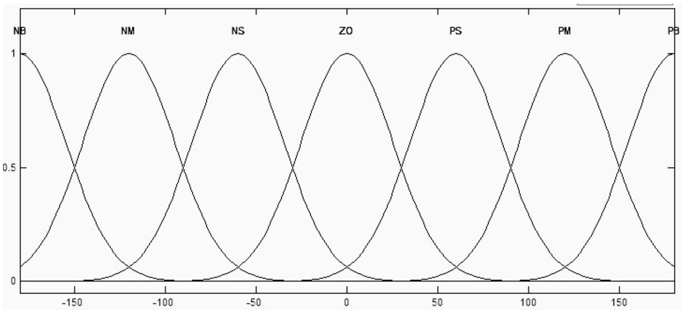


Fig. 4. Member function of φ

Table 1. Fuzzy control rules

φ_c	φ						
	NB	NM	NS	ZE	PS	PM	PB
NB	BZS	BSM	MMZ	MBS	SMZ	SSM	ZZS
NM	BZM	MSM	MBS	SBZ	SBS	ZSM	ZZM
NS	BZB	MZB	SMS	ZBS	SMS	SZB	BZB
ZE	BZB	SZB	ZBS	ZBZ	ZBS	SZB	BZB
PS	BZB	SZB	SBS	ZBS	SBS	MZB	BZB
PM	ZZM	ZSM	SMS	SBZ	MMS	MSM	BZM
PB	ZZS	ZSM	SMZ	MBZ	MMZ	BZM	BZS

4 System Simulation and Analysis

According to the simulation model which established above, simulation analysis is carried on, the transfer function of the simulation object is:

$$G(s) = \frac{4}{s(s+2.5)} \tag{2}$$

Add step signal 1 to fuzzy control system, through adjusting the PID parameters starting value many times and revises the fuzzy rule repeatedly, satisfying response curves are obtained, Figure 5 is traditional PID and the fuzzy PID controller system's step response curve. It can be find that, the fuzzy PID function output surpasses the traditional PID function greatly, and non-overshoot control has realized.

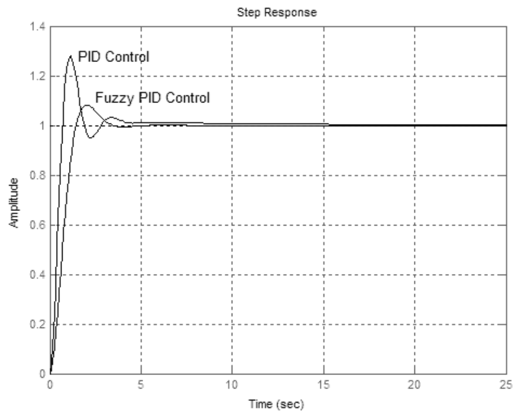


Fig. 5. Step response curves

5 Conclusion

Based on the analysis of LUAVPS pan-tilt, a pan-tilt fuzzy PID controller is designed combining traditional PID control theory and fuzzy theory. Simulation experiment result shows that this PID controller can achieve no overshoot control, and maintain traditional PID controller advantage; Especially has a strong adaptability and good robustness when object parameters or structure changes.

Acknowledgements. The work was supported by The key project of institutes serving for the economic zone on the western coast of the Taiwan strait (2009(8)); The key project of education department of Fujian province, PRC(JK2009017); The National Natural Science Foundation of Xiamen City, PRC (3502Z2093018).

References

1. Jin-kun, L.: Intelligent Control. Publishing House of Electronics Industry, Beijing (2009)
2. Dong-lin, Z., Ju-hua, H., Qun, Z.: Application of fuzzy self-tuning PID controller in electric power steering system. Control industry automation 31(6), 72–75 (2009)
3. Li-da, L., Bei-zhan, W., Yun-qi, Z.: Research and application of adaptable fuzzy-PID controller. Construction machine 8, 59–62 (2009)
4. Yong-hua, T.: New PID controller and application. Machine Press, Beijing (2002)
5. Yi-hui, J.: Process control. Tsinghua University Press, Beijing (1993)
6. Yong-jun, Z., Hua-dong, S.: Design and Simulation of A Fuzzy PID Controller Based on Matlab. Control & Automation 25(1-1), 48–50 (2009)
7. Jian-min, Z., Tao, W., Zhong-li, W.: Intelligent control principle and application. Metallurgy Industry Press (2003)
8. Han-sheng, Z.: Autopilot design Based on Fuzzy PID. China New Technolo-gies and Products 1, 2–3 (2009)
9. Chen, S.-l., et al.: Fuzzy Sets Logic and its Applications. Science Press (2005)

ω-Convergence Theory of Filters in ω-Molecular Lattices

Qiang Ren

Sanming University, Sanming, Fujian 365004, P.R. China
e-mail: fjsmrg@126.com

Abstract. In this paper, an ω-convergence theory of filters in an ω-molecular lattice is established. By means of the ω-convergence theory, some important characterizations with respective to the ω-closed sets and (ω_1, ω_2)-continuous generalized order-homomorphisms are obtained. Moreover, the mutual relationships among ω-convergence of molecular nets, ω-convergence of ideals and ω-convergence of filters are given in ω-molecular lattices.

Keywords: Fuzzy lattice, ω-molecular lattice, filter, ideal, fuzzy mapping, ω-convergence.

1 Introduction

The Moore-Smith convergence theory in topological molecular lattices was first introduced by Wang [1]. Since then, many convergence theories, such as θ-convergence theory [2], δ-convergence theory [9], N-convergence theory [3], SR-convergence theory [4], σ-convergence theory [5], and so on [2-5,9], were presented by means of multifarious closure operators. In order to unify various convergence theories, a generalized molecular lattice which called the ω-molecular lattices was introduced [6-8]. In this paper, an ω-convergence theory of filters in ω-molecular lattices.

Throughout the paper, L denotes a fuzzy lattice while M denotes the set consisting of all molecules [1], i. e., nonzero $\vee$-irreducible elements in L. 0 and 1 are the least and the greatest element of L respectively.

Definition 1.1. [1] *Let L be a complete lattice.* $e\in L$, $B\subset L$. *B is called a minimal family of e if* $B\neq\varnothing$ *and*

(i) $\sup B= e$;
(ii) $\forall A\subset L$, $\sup A\geq e$ *implies that* $\forall x\in B$, *there exists a* $y\in A$ *such that* $y\geq x$.

According to Hutton [19], in a completely distributive lattice, each element $e\in L$ has a greatest minimal family which will be denoted by $\beta(e)$. For each $e\in M$, $\beta^*(e)=\beta(e)\cap M$ is a minimal family of e and is said to be the standard minimal family of e.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 561–568.
springerlink.com

Definition 1.2. [6] *Let L be a molecular lattice.*

(i) An operator ω: $L\to L$ is said to be an ω-operator if (1) $\forall A, B\in L$ and $A\leq B$; (2) $\forall A\in L$, $A\leq \omega(A)$.
(ii) An element $A\in L$ is called an ω-set if $\omega(A)=A$.
(iii) Put $\Omega=\{A\in L / \omega(A)=A\}$, and call the pair (L, Ω) an ω-molecular lattice.

Definition 1.3. [6] *Let (L, Ω) be an ω- molecular lattice, $A\in L$ and $e\in M$. If there exists $Q\in\Omega$ such that $e\not\leq Q$ and $P\leq Q$, then call P an ωR-neighborhood of e. The collection of all ωR-neighborhoods of e is denoted by $\omega\eta(e)$.*

Definition 1.4. [6] *Let (L, Ω) be an ω-molecular lattice, $A\in L$ and $e\in M$. If $A\not\leq P$ for each $P\in\omega\eta(e)$, then α is said to be an ω-adherence point of A, and the union of all ω-adherence points of A is called the ω-closure of A, and denoted by $\omega cl(A)$. If $A=\omega cl(A)$, then call A an ω-closed element. If A is an ω-closed element, then we say that A' is an ω-open element. If $P=\omega cl(P)$ and $e\not\leq P$, then P is said to be an ω-closed R-neighborhood (briefly, ωCR-neighborhood) of e, and the collection of all ωCR-neighborhoods of e is denoted by $\omega\eta^{-}(e)$.*

Definition 1.5. [6] *Let (L,Ω) be an ω-molecular lattice, $A\in L$ and $\omega int(A)=\bigvee\{B\in L / B\leq A$ and B is an ω-open element in L $\}$. We call $\omega int(A)$ the ω-interior of A. Obviously, A is ω-open if and only if $A=\omega int(A)$.*

Proposition 1.1 [23] *Let L be a completely distributive lattice. Then each element of L is a union of $\bigvee$-irreducible elements.*

2 ω-Convergence of Filters

In this section, we shall present the concepts of ωQ-neighborhoods (resp. ωOQ-neighborhoods) of a molecule and ω-convergence of a filter in an ω-molecular lattice, and discuss their properties.

Definition 2.1. *Let (L, Ω) be an ω-molecular lattice, $B\in L$ and $e\in M$. If there is an ω-open element G such that $e\not\leq G'$ and $G\leq B$, then we say that B (resp. G) is an ωQ-neighborhood (resp. ωOQ-neighborhood) of e, and the collection of all ωQ-neighborhoods (resp. ωOQ-neighborhoods) of e is denoted by $\omega\mu(e)$ (resp. $\omega\mu^{o}(e)$).*

Evidently, every ωQ-neighborhood (resp. ωOQ-neighborhood) of e is a Q-neighborhood (resp. open Q-neighborhood) of e when ω is the fuzzy closure operator, and B (resp. G) is an ωQ-neighborhood (resp. ωOQ-neighborhood) of e if and only if B' (resp. G') is an ωR-neighborhood (resp. ωCR-neighborhood) of e.

Definition 2.2. *Let (L, Ω) be an ω-molecular lattice, e∈M and let F be a filter in L. Then:*

(i) e is said to be an ω-limit point of F, or F ω-converges to e, in symbols, F→ωe, if ωμo(e)⊂F. The union of all ω-limit points of F will be denoted by ω-limF.
(ii) e is said to be an ω-cluster point of F, or F ω-accumulates to e, in symbols,F∝ωe, if F⋀ G≠0 for each G∈ωμo(e) and each F∈F. The union of all ω-cluster points of F will be denoted by ω-adF.

Theorem 2.1. Let (L, Ω) be anω-molecular lattice, e∈ M and let F be a filter in L. Then:

(1) $F\to_\omega e$ if and only if $F\to_\omega b$ for each $b\in\beta^*(e)$;
(2) $F\propto_\omega e$ if and only if $F\propto_\omega b$ for each $b\in\beta^*(e)$;
(3) $F\propto_\omega e$ if and only if $e\le\omega\mathrm{cl}(F)$ for each $F\in$F.

*Proof***.** (1) Suppose that $F\to_\omega e$, $b\in\beta^*(e)$ and $G\in\omega\mu^o(b)$. Then $G\in\omega\mu^o(e)$ because of the fact that $b\not\le G'$ and $b\le e$, and hence $G\in$F by $F\to_\omega e$. Conversely, if e is not an ω-limit point of F, then there exists $G\in\omega\mu^o(e)$ such that $G\notin$ F. Since $e=\beta^*(e)$, there is $b\in\beta^*(e)$ with $G\in\omega\mu^o(b)$. This means that b is not an ω-limit point of F. Hence, the sufficiency is proved.

(2) Similar to the proof of (1).

(3) Let $F\propto_\omega e$. Then $F\wedge G\neq 0$ for each $G\in\omega\mu^o(e)$ and each $F\in$F by Definition 2.2(2), equivalently, $F\not\le G'$ for each $G'\in\omega\eta^-(e)$ and each $F\in$F. Therefore, $e\le\omega\mathrm{cl}(F)$ for each $F\in$F. Conversely, if $e\le\omega\mathrm{cl}(F)$ for each $F\in$F, then $F\not\le G'$ for each $G'\in\omega\eta^-(e)$ by Definition 2.4., in other words, $F\wedge G\neq 0$ for each $G\in\omega\mu^o(e)$. Consequently, $F\propto_\omega e$ by arbitrariness of F in F.

Proposition 2.1. *Let (L, Ω) be an ω-molecular lattice, b, d∈M, and let F be a filter in L. Then:*

(1) if $F\to_\omega d$ and $b\le d$, then $F\to_\omega b$;
(2) if $F\propto_\omega d$ and $b\le d$, then $F\propto_\omega b$.

Proof. (1) Let $F\to_\omega d$ and $b\le d$. Then $G\in\omega\mu^o(d)$ for each $G\in\omega\mu^o(b)$, and thus $G\in$F. This implies that $\omega\mu^o(b)\subset$F, hence $F\to_\omega b$.

(2) Similar to the proof of (1).

Theorem 3.2. *Let* (L^X, Ω) *be an* ω-molecular lattice, *e*∈*M, and let* F *be a filter in L. Then*:

(1) $F\to_\omega e$ *if and only if* $e\le\omega$-limF;
(2) $F\propto_\omega e$ *if and only if* $e\le\omega$-adF;
(3) ω-limF $\le\omega$-adF.

Proof. (1) If $\mathrm{F}\to_{\omega}e$, then $e\leq\omega$-limF by the definition of ω-limF. Conversely, if $e\leq\omega$-limF, then for each $b\in\beta^*(e)$, there exists an ω-limit point d of F with $b\leq d$ by virtue of the fact that $e=\sup\beta^*(e)$ and the definition of ω-limF. Consequently, $\mathrm{F}\to_{\omega}e$ according to Proposition 2.1 and Theorem 2.1.

(2) Similar to the proof of (1).

(3) Let $e\leq\omega$-limF. Then $\mathrm{F}\to_{\omega}e$ by (1). In accordance with Definition 2.2(i), we know that $G\in\mathrm{F}$ for each $G\in\omega\mu^{\mathrm{o}}(e)$. Therefore, $F\wedge G\neq 0$ for each $G\in\omega\mu^{\mathrm{o}}(e)$ and each $F\in\mathrm{F}$ by the definition of filter, thus $e\leq\omega$-adF in the light of Definition 2.2(ii) and (2).

Proposition 2.2. *Suppose that* F_1 *and* F_2 *are two filters in an* ω-molecular lattice (L,Ω), *which* F_2 *is finer than* F_1 (*i.e.*, $\mathrm{F}_1\subset\mathrm{F}_2$) *and* $e\in M$. *Then*:

(1) *if e is an* ω*-limit point of* F_1, *then e is also an* ω*-limit point of* F_2;
(2) *if e is an* ω*-cluster point of* F_2, *then e is also an* ω*-cluster point of* F_1.

Proof. (1) If e is an ω-limit point of F_1, then $\mu^{\mathrm{o}}(e)\subset\mathrm{F}_1$. Since $\mathrm{F}_1\subset\mathrm{F}_2$, $\mu^{\mathrm{o}}(e)\subset\mathrm{F}_2$. Therefore, e is an ω-limit point of F_2.

(2) Let F_2 ω-accumulates to e. Then for each $G\in\omega\mu^{\mathrm{o}}(e)$ and each $F\in\mathrm{F}_2$ we have $F\wedge G\neq 0$, specially, for each $F\in\mathrm{F}_1$, $F\wedge G\neq 0$ by virtue of $\mathrm{F}_1\subset\mathrm{F}_2$. Hence, F_1 ω-accumulates to e.

Theorem 2.3. *Let* F *be a filter in an* $L\omega$*-space* (L,Ω) *and* $e\in M$. *Then e is an* ω*-cluster point of* F *if and only if there exists a filter* F* *which is finer than* F *such that e is an* ω*-limit point of* F*.

Proof. Assume that e is an ω-cluster point of F. By Definition 2.2(ii), $F\wedge G\neq 0$ for each $G\in\omega\mu^{\mathrm{o}}(e)$ and each $F\in\mathrm{F}$. Write $\mathrm{F}^*=\{H\in L^X \mid F\wedge G\leq H$ for each $G\in\omega\mu^{\mathrm{o}}(e)$ and each $F\in\mathrm{F}\}$; then F* is a filter which is finer than F, and $G\in\mathrm{F}^*$ for each $G\in\omega\mu^{\mathrm{o}}(e)$. This implies that e is an ω-limit point of F*.

Conversely, suppose that F* is a filter which is finer than F, and F* ω-converges to e. According to Definition 2.2(i), for each $G\in\omega\mu^{\mathrm{o}}(e)$ we have $G\in\mathrm{F}^*$, hence $F\wedge G\in\mathrm{F}^*$ for each $F\in\mathrm{F}^*$ and each $G\in\omega\mu^{\mathrm{o}}(e)$, and hence $F\wedge G\neq 0$ by the definition of filter. This shows that e is an ω-cluster point of F.

Theorem 2.4. *Let* (L,Ω) *be an* ω-molecular lattice, $A\in L$ *and* $e\in M$. *Then the following conditions are equivalent*:

(1) *e is an* ω*-adherence point of A*;
(2) *there exists a filter* F *with* $A\in\mathrm{F}$ *such that e is an* ω*-limit point of* F;
(3) *there exists a filter* F *with* $A\in\mathrm{F}$ *such that e is an* ω*-cluster point of* F.

Proof. (1)⇒(2): Suppose that e is an ω-adherence point of A; then $A\wedge G\neq 0$ for each $G\in\omega\mu^{\mathrm{o}}(e)$. Let

$$\mathrm{F} = \{F \in L \mid \text{there exists } G \in \omega\mu^{\circ}(e) \text{ with } A \wedge G \le F\}.$$

One can easily see that F is a filter with $A \in \mathrm{F}$ and $\omega\mu^{\circ}(e) \subset \mathrm{F}$, i.e., e is an ω-limit point of F.

(2)⇒(3): If there exists a filter F with $A \in \mathrm{F}$ such that e is an ω-limit point of F, then $A \wedge G \neq 0$ by $A \wedge G \in \mathrm{F}$ for each $G \in \omega\mu^{\circ}(e)$, and thus e is an ω-cluster point of F.

(3)⇒(1): If there exists a filter F with $A \in \mathrm{F}$ such that e is an ω-cluster point of F, then $A \wedge G \neq 0$ for each $G \in \omega\mu^{\circ}(e)$, in other words, $A \not\leq G'$ for each $G' \in \omega\eta^{-}(e)$. Consequently, e is an ω-adherence point of A.

Theorem 2.5. *Let* (L, Ω) *be an* $L\omega$*-space and* $A \in L$. *Then the following conditions are equivalent*:

(1) *A is an ω-closed element*;
(2) *for each filter* F *containing A as an element in L*, ω-limF $\le A$;
(3) *for each filter* F *containing A as an element in L*, ω-adF $\le A$.

Proof. (1)⇒(2): Suppose that A is an ω-closed element, F is a filter containing A as an element, and $e \in M$. If $e \le \omega$-limF. Then $e \le \omega\mathrm{cl}(A) = A$ in line with Theorem 2.4. Therefore, ω-limF $\le A$.

(2)⇒(3): It follows from (2) and Theorem 2.2(3).

(3)⇒(1): Assume that ω-adF $\le A$ for each filter F containing A as an element and $e \le \omega\mathrm{cl}(A)$. Then by Theorem 2.4 we know that $e \le \omega$-adF $\le A$. This means that $\omega\mathrm{cl}(A) \le A$, i.e., A is ω-closed.

Theorem 2.6. *If* F *be a filter in an* ω-molecular lattice (L, Ω), *then* ω-limF *and* ω-adF *are both ω-closed sets in* (L, Ω).

Proof. For each $e \in M$, if $e \le \omega\mathrm{cl}(\omega\text{-limF})$, then ω-limF $\not\leq G$ for each $G' \in \omega\mu^{\circ}(e)$,. With reference to Proposition 1.1 we can choose a molecule $b \le \omega$-limF such that $b \not\leq G$, thus $G' \in \omega\mu^{\circ}(b)$. Consequently, $G' \in \mathrm{F}$ by $b \le \omega$-limF. This shows that $e \le \omega$-limF, and thus $\omega\mathrm{cl}(\omega\text{-limF}) \le \omega$-limF. On the other hand, ω-limF $\le \omega\mathrm{cl}(\omega\text{-limF})$ by Theorem 2.5 in [6].

Similarly, we can easily verify that ω-adF is also an ω-closed element in (L, Ω).

Definition 2.3. *Let (L, Ω) be an ω-molecular lattice, $e \in M$, and let F_0 be a filter base in L. Then:*

(i) e is said to be an ω-limit point of F_0, or F_0 ω-converges to e, in symbols, $F_0 \to_{\omega} e$, if $F \to_{\omega} e$ where F is the filter generated by F_0, i.e.,

$$F = \{F \in L \mid \text{there exists } H \in F_0 \text{ such that } H \le F\}.$$

The union of all ω-limit points of F_0 will be denoted by ω-limF_0.

(ii) e is said to be an ω-cluster point of F_0, or F_0 ω-accumulates to e, in symbols, $F_0 \propto_\omega e$, if $F \propto_\omega e$ where F is the filter generated by F_0. The union of all ω-cluster points of F_0 will be denoted by ω-adF_0.

Theorem 2.7. *Let* F_0 *be a filter base in an* ω-molecular lattice e (L, Ω) *and* $e \in M$. *Then*:

(1) $\mathrm{F}_0 \rightarrow_\omega e$ *if and only if every* ωOQ*-neighborhood of e contains a member of* F_0;
(2) $\mathrm{F}_0 \propto_\omega e$ *if and only if every* ωOQ*-neighborhood of e intersects all member of* F_0.

Proof. It follows straightforward from Definition 2.2 and Definition 2.3.

3 Some Applications of ω-Convergence of Filters

In this section, we shall give some characterizations of (ω_1, ω_2)-continuous order-homomorphisms and ω-separations by means of ω-convergence theory of filters.

Definition 3.1. [7] *Let f be an order-homomorphism from an* ω_1*-molecular lattice* (L_1, Ω_1) *into an* ω_2*-molecular lattice* (L_2, Ω_2). *Then:*

(i) f is called (ω_1, ω_2)*-continuous if* $f^{-1}(B) \in \omega_1 O(L_1)$ *for each* $B \in \omega_2 O(L_2)$;
(ii) f is called (ω_1, ω_2)*-continuous at* $e \in M$ *if* $f^{-1}(B) \in \omega_1 \eta^{-}(e)$ *for each* $B \in \omega_2 \eta^{-}(f(e))$.

Theorem 3.1. [7] *Let f be an* order-homomorphism *from an* ω_1-molecular lattice (L_1, Ω_1) *into an* ω_2-molecular lattice (L_2, Ω_2). *Then f is* (ω_1, ω_2)*-continuous if and only if for every* $e \in M$, *f is* (ω_1, ω_2)*-continuous at e.*

Theorem 3.2. *Let f be an* order-homomorphism *from an* ω_1- molecular lattice (L_1, Ω_1) *into an* ω_2- molecular lattice (L_2, Ω_2), *and* $e \in M$. *Then f is* (ω_1, ω_2)*-continuous at e if and only if* $f^{-1}(B) \in \omega_1 \mu^{o}(e)$ *for each* $B \in \omega_2 \mu^{o}(f(e))$.

Proof. Since $B \in \omega_2 \mu^{o}(f(e))$ if and only if $B' \in \omega_2 \eta^{-}(f(e))$, and $f^{-1}(B) \in \omega_1 \mu^{o}(e)$ if and only if $f^{-1}(B') \in \omega_1 \eta^{-}(e)$, the proof is obvious by Definition 3.1(ii).

Theorem 3.3. *Let f be an* order-homomorphism *from an* ω_1- molecular lattice (L, Ω_1) *into an* ω_2- molecular lattice (L_2, Ω_2). *Then f is* (ω_1, ω_2)*-continuous if and only if for every filter base* F_0 *in* (L_1, Ω_1) *and the filter base* $f(\mathrm{F}_0)=\{f(F) \mid F \in \mathrm{F}_0\}$ *in* (L_2, Ω_2) *we have* $f(\omega_1\text{-lim}\mathrm{F}_0) \le \omega_2\text{-lim} f(\mathrm{F}_0)$.

Proof. Suppose that f is (ω_1, ω_2)-continuous, F_0 is an filter base in (L_1, Ω_1) and $f(\mathrm{F}_0)=\{f(F) \mid F \in \mathrm{F}_0\}$ is the filter base in (L_2, Ω_2). Then for each molecule $d \le f(\omega_1\text{-lim}\mathrm{F}_0)$, there exists a molecule $e \le \omega\text{-lim}\mathrm{F}_0$ with $d=f(e)$. We affirm that $d \le \omega_2\text{-lim}f$

(F_0). In fact, for each $B\in \omega_2\eta^-(f(e))$ we have $f^{-1}(B)\in \omega_1\eta^-(e)$ being the continuity of f, in other words, for each $B'\in \omega_2\mu^o(f(e))$ we have $f^{-1}(B')\in \omega_1\mu^o(e)$. Hence, there exists a member $F\in F_0$ such that $F\leq f^{-1}(B')$, i.e., $f(F)\leq B'$ according to Theorem 2.7. This means that $d\leq\omega_2$-lim$f(F_0)$.

Conversely, assume that the condition of the theorem holds. If f is not (ω_1,ω_2)-continuous, then there exists an ω_2-closed element B in (L_2, Ω_2) such that $f^{-1}(B)\neq\omega_1\mathrm{cl}(f^{-1}(B))$, i.e., there is a molecule $e\leq\omega_2\mathrm{cl}(f^{-1}(B))$ with $e\not\leq f^{-1}(B)$ by Proposition 1.1. Hence, we can choose a filter base F_0 in (L^X, Ω_1) which contains $f^{-1}(B)$ as member such that e is an ω-limit point of F_0 in the light of Theorem 2.7, and thus $f(e)\leq f(\omega_1$-lim$F_0)\leq\omega_2$-lim$f(F_0)$ according to the assumption. Since B is ω_2-closed, $f(e)\leq B$ according to Theorem 2.5. This contradicts $e\not\leq f^{-1}(B)$. Therefore, f is (ω_1,ω_2)-continuous.

Theorem 3.4. *An $L\omega$-space (L, Ω) is an ωT_2 space if and only if every filter in (L, Ω) has at most one ω-limit point..*

Proof. Assume that (L, Ω) is an ωT_2 space and F is a filter in (L, Ω). If e and d are two disjoint ω-limit points of F, then there exist $G\in \omega\eta^-(e)$ and $H\in \omega\eta^-(d)$ such that $G\vee H=1$, in other words, there exist $G'\in \omega\mu^o(e)$ and $H'\in \omega\mu^o(d)$ such that $G'\wedge H'=0$. Hence, $G'\wedge H'\notin$ F, i.e., either $G'\notin$ F or $H'\notin$ F. This contradicts that e and d are both ω-limit points of F. Therefore, the necessity is proved.

Conversely, if (L, Ω) is not ωT_2, then there are $e, d\in M$ with $e\neq d$, such that for each $G\in \omega\mu^o(e)$ and each $H\in \omega\mu^o(d)$ satisfying $G'\vee H'\neq 1$, equivalently, $G\wedge H\neq 0$. Let $F=\{F\in L\,|\,G\wedge H\leq F,\ G\in \omega\mu^o(e) \text{ and } H\in \omega\mu^o(d)\}$. Obviously, $\omega\mu^o(e)\subset$ F and $\omega\mu^o(d)\subset$ F, thus e and d are disjoint ω-limit points of F. This implies that the sufficiency holds.

Acknowledgements. The paper is supported by the Natural Science Foundation of Fujian Province of China (A0810014).

References

1. Wang, G.J.: Theory of topological molecular lattices. Fuzzy Sets and Systems 47, 351–376 (1992)
2. Chen, S.L.: θ-convergence theory of ideals on topological molecular lat-tices. J. Math. 13(2), 163–169 (1993)
3. Chen, S.L., Chen, S.T.: A new extension of fuzzy convergence. Fuzzy Sets and Systems 109, 1199–1204 (2000)
4. Chen, S.L., Wu, J.R.: SR-convergence theory in fuzzy lattices. Information Sciences 125, 233–247 (2000)
5. Chen, S.L., Chen, S.T., Wang, X.G.: σ-convergence theory and its applica-tions in fuzzy lattices. Information Sciences 165(1), 45–58 (2004)

6. Chen, S.L.: On ω-molecular lattices. In: Proceeding of ICMLC 2005, Guang-zhou, vol. I, pp. 2619–2623 (2005)
7. Chen, S.L.: Separation axioms in ω-molecular lattices. In: Proceeding of ICMLC 2006, Dalian, vol. I, pp. 1801–1804 (2006)
8. Chen, S.L.: ω-convergence theory of molecular nets in ω-molecular lattices. In: Proceeding of ICNC-FSKD 2010 Conference, vol. I (2010)
9. Cheng, J.S.: Some properties of δ-continuous order-homomorphisms. Fuzzy Systems and Math. 11(4), 38–41 (1997)
10. Hutton, B.: Uniformities on fuzzy topological spaces. J. Math. Anal. Appl. 58, 559–571 (1977)

The Structures and Constructions of Power Group

Yu-bin Zhong[1] and Hong-xing Li[2,3,⋆]

[1] School of Mathematics and Information Sciences, Guangzhou University, Guangzhou 510006, China
e-mail: Zhong_yb@163.com

[2] School of Mathematical Science, Beijing Normal University, Beijing 100875, China

[3] School of Electronic and Information Engineering, Dalian University of Technology, Dalian 116024, China
e-mail: lhxqx@bnu.edu.cn

Abstract. With the applications of topological upgrading [1], some problems of upgrading mathematical structure to its power set have caused wide concerns. This paper studies the upgrading from group structure to its power set and gives the definition, examples and ways of construction of power groups, then shows the relationship between power groups and common groups and characterized power groups by common groups, and obtains a series of constructive conclusions and achieves many breakthroughs of upgrading a group to its power set.

Keywords: Upgrade, power group, structure, homomorphism, isomorphism.

1 Introduction

Professor Wang Peizhuang [1] raised the problem of topological upgrading in 1981 and obtained a series of systematic theoretical results and applications [2-9]. Hypergroup which is raised by H.S.Wall [10] has important applications in theoretical physics including quantum mechanics and string theory [11,12], thus this kind of upgrading of mathematical structure has caused wide concern over the recent years [13-21]. This paper studies the upgrading from group structure to its power set. By applying the principle of extension and set valued mapping to classic algebra, it obtains a new kind of algebraic structure named power group, gives several examples and the ways of construction of power groups, and shows different kinds of relationship between

⋆ Corresponding author.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 569–579.
springerlink.com

power groups and common groups. Finally, in the last part of this paper we studies the general structure, quasi-order structure, homomorphism and isomorphism of power groups, obtains a series of constructive conclusions and achieves many breakthroughs of upgrading a group to its power set. This research will be introduced concretely and the conclusions will be presented in detail in the following.

2 Definitions and Constructions of Power Group

Here in this paper we always assemble that G denotes a group, let $\mathcal{P}_0(G) \triangleq \mathcal{P}(G)\backslash\{\phi\}$,$\mathcal{P}(G)$ denotes the power set of G. Now we define an algebraic operation on $\mathcal{P}_0(G)$:

$$AB \triangleq \{ab|a \in A, b \in B\}. \tag{1}$$

Obviously $\mathcal{P}_0(G)$ forms a semigroup with respect to composition (1), and it contains e which is the identity of G.

Definition 2.1. *Let $\mathcal{G} \subset \mathcal{P}_0(G)$ be a non-empty class of sets. If $\mathcal{G}$ forms a group with respect to the operation defined by (1), then we say $\mathcal{G}$ is a power group of G whose identity is E . Especially, we say $\mathcal{G}$ is a regular power group of G iff $e \in E$.*

Notice that $E^2 = E$, so it's easy to see that E is a subgroup of G. What's more, the quotient group is a particular case of power group, its identity is a normal group of group G.

Theorem 2.1. *If $\mathcal{G}$ is a power group of G,$\forall A \in \mathcal{G}$, denote $|A|$ to be the number of elements which A contains, then*

1)$(\forall A \in \mathcal{G})(|A| = |E|)$;

2)$(\forall A, B \in \mathcal{G})(A \cap B \neq \phi \Rightarrow |A \cap B| = |E|)$.

Proof. 1) On one hand,$AE = A \Rightarrow (\forall a \in A)(aE \subset AE = A) \Rightarrow |E| = |aE| \leq A$;on the other hand, $A^{-1}A = E \Rightarrow (\forall b \in A^{-1})(bA \subset A^{-1}A = E) \Rightarrow |A| = |bA| \leq |E|)$.

2) First,$|A \cap B| \leq |A| = |E|$;then $c \in A \cap B \Rightarrow cE \subset A \cap B \Rightarrow |E| = |cE| \leq |A \cap B|$.

Considering that E is a subgroup of G, this question comes naturally as follows

Is it possible to construct a power group of G using a subsemigroup S of group G when $S = E$? We studied this problem and got an affirmative answer, which we list as follows.

Suppose S is a subgroup of G,$\forall a, b \in G$,if $\exists h \in S$,s.t. $a = bh$, then we say a is congruent to b right-semi-modulo S, denoted by

$$a \doteq b(r - semodS). \tag{2}$$

It's easy to verify that $\doteq$ is a transitive relation, what's more,$\doteq$ satisfies self-opposite (therefore $\doteq$is a similar relation) iff $e \in S$. Now for $\forall a \in G$, let

$$aS \triangleq \{b \in G | b \doteq a(r - semodS)\}, \tag{3}$$

it's obvious that $aS = \{ah | h \in S\}$. We call aS to be the left quasi-coset of S. Similarly, we can get the concept of right quasi-coset of S, denoted by Sa. We call Sa a quasi-coset iff $\forall a \in G, Sa = aS$.

Remark,$a \in aS$ is not valid generously. For example, set G to be the addictive group of real numbers, make $S = (0, +\infty)$, then S is a subsemigroup of G, but $\forall a \in G, a \notin a+S = (a, +\infty)$, if strengthen the constraint conditions of S, then we have this conclusion:$e \in S \Rightarrow (\forall a \in G)(a \in aS)$;otherwise $(\exists a \in G)(a \in aS) \Rightarrow e \in S$.

Theorem 2.2. *Suppose H is a subgroup of G,$S \subset G$ which satisfies $S^2 = S$.If S is a quasi-coset of H, which means $(\forall a \in H) aS = Sa$, then $\mathcal{G} \triangleq \{aS | a \in H\}$ is a power group of G and $E = S$.*

Proof. Define a mapping as $f : H \to \mathcal{G}, a \mapsto aS$,it's obvious that f is a surjection. Notice that $f(ab) = (ab)S = (ab)S^2 = a(bS)S = a(Sb)S = (aS)(bS) = f(a)f(b)$,so H and $\mathcal{G}$ are homomorphic which means $H \sim \mathcal{G}$, thus $\mathcal{G}$ is a group. In addition, from $f(e) = eS = S$, we get $S = E$.

This theorem shows that once a certain sort of subsemigroup S of G is given, we can construct a power group $\mathcal{G}$ on G using a of subsemigroup H of G , and $S = E$. What's more, we have $H \sim \mathcal{G}$. Thus we solve the problem given previously.

Remark 1: For $H \sim \mathcal{G}$,so $H/kerf \cong \mathcal{G}$, this means that the power group $\mathcal{G}$ formed in this way must be isomorphic to a quotient group of a certain subgroup of G.

Remark 2: It's easy to see that $S^2 = S \Rightarrow S$ is a semigroup, but the inverse proposition is not valid. For example, set G to be the addictive group $(\mathbb{R}, +)$, $S = [1, +\infty)$,then S is a subsemigroup of G,but $S+S = [2, +\infty) \neq S$. But if we let $e \in S$, then $S^2 = S \Leftrightarrow S$ is a subsemigroup of G.

Example 2.1. Let G be the multiplicative group of positive real numbers. Make $E = [1, +\infty)$ and H to be the set of all the rational numbers in G , they satisfy condition 2, thus $\mathcal{G} = \{aE | a \in H\}$ is a power group of G . It's obvious that $\mathcal{G} = \{[a, +\infty) | a \in H\}$. Let $f : H \to \mathcal{G}, a \mapsto [a, +\infty)$,then $H \cong \mathcal{G}$.

Example 2.2. Let G be the addictive group of all real numbers. Make $E = (0, +\infty)$ and H to be the set of all integers, then $\mathcal{G} = \{n + E | n \in H\}$ is a power group of G, and $H \cong \mathcal{G}$. It's worth noting that $0 \notin E$, and the elements of $\mathcal{G}$ forms a countable chain:

$$\cdots \supset (-2, +\infty) \supset (-1, +\infty) \supset E \supset (1, +\infty) \supset (2, +\infty) \supset \cdots.$$

Example 2.3. Let $(G,+,\geq)$ be a addictive group with semi order, and $\mathcal{G} \triangleq \{[a,b]|a,b \in G\}$ in which $[a,b] \triangleq \{c \in G|a \geq c \geq b\}$.Define an algebraic operation "+" on $\mathcal{G}$ as:

$$[a_1,b_1]+[a_2,b_2] \triangleq [a_1+a_2,b_1+b_2]$$

it's easy to see that $\mathcal{G}$ is a power group of G,and $E=[0] \triangleq \{0\}$. In addition, the mapping$f:\mathcal{G} \to G,[a,b] \mapsto a, h:\mathcal{G} \to G,[a,b] \mapsto b$ are both epimorphism. Obviously, $ker\ f=\{[0,b]|b \in G\}, ker\ h=\{[a,0]|a \in G\}$, therefore $\mathcal{G}/ker\ f \cong G \cong ker\ f \cong ker\ h \cong \mathcal{G}/ker\ f.$

3 Relationship between Power Groups and Common Groups

This section mainly studies the relationship between power groups and common groups. To simplify the statement, we first employ some notions and brief introductions:

Definition 3.1. *Let $\mathcal{G}$ be a power group of G . Denote $G^* \triangleq \bigcup\{A|A \in \mathcal{G}\}, G^0 \triangleq \{a \in G^*|a^{-1} \in G^*\}$.*

Theorem 3.1. 1) G^* is a subsemigroup of G;
2) $G^0 \neq \varnothing \Leftrightarrow e \in G^0$;
3) $G^0 \neq \varnothing \Leftrightarrow G^0$ *is a subsemigroup of* G

Proof. 1)$a,b \in G^* \Rightarrow (\exists A,B \in \mathcal{G})(a \in A, b \in B) \Rightarrow ab \in AB \subset G^*$.
2) The conclusion is obvious.
3) On one hand, from G^0 is a subsemigroup of G it's easy to see that $G^0 \neq \varnothing$; on another hand, $a,b \in G^0 \Rightarrow a^{-1},b^{-1} \in G^* \Rightarrow (ab^{-1})^{-1}=ba^{-1} \in G^* \Rightarrow G^0$ is a subsemigroup of G.

Theorem 3.2. *Let $\mathcal{G}$ be a regular power group of G. If $G^0 \subset H \subset G^*$, then H is a subgroup of $G \Leftrightarrow H=G^0$.*

Proof. Suppose H is a subgroup of G, if $H \neq G^0$,let $a \in H \backslash G^0$, then $a^{-1} \in H \subset G^*$, so $a \in G^0$, contradiction occurred, so $H=G^0$. On another hand, if $H=G^0$, then from theorem 3.1(3) we know H is a subgroup of G.

Definition 3.2. *Suppose $\mathcal{G}$ is a power group of G,A^{-1} is the inverse of A,$A^{\odot}=\{x^{-1}|x \in A\}$ is called the inverse set of A,if $\forall A \in \mathcal{G}, A^{-1}=A^{\odot}$, then we say $\mathcal{G}$ is a power group whose inverse and inverse set are uniform, which can be called 'uniform power group' for short.*

Theorem 3.3. *If power group $\mathcal{G}$ is uniform, then $G^* \triangleq \bigcup\{A|A \in \mathcal{G}\}$ is a subgroup of G.*

Proof. $\forall a,b \in G^*, \exists A,B \in \mathcal{G}$,s.t.$a \in A, b \in B$.So $ab^{-1} \in AA^{\odot}=AB^{-1} \in \mathcal{G}$,that is $ab^{-1} \in G^*$,thus G^* forms a subgroup of G.

Theorem 3.4. *$\mathcal{G}$ is a uniform power group $\Leftrightarrow$ the identity E is a subgroup of G.*

Proof. " $\Rightarrow$ " $\mathcal{G}$ is uniform $\Rightarrow E^{-1} = E^{\odot} \Rightarrow (\forall a, b \in E)(ab^{-1} \in E \bullet E^{\odot} = E \bullet E^{-1} = E) \Rightarrow E \leq G$." $\Leftarrow$ "$E \leq G \Rightarrow (\forall a \in A)(aE = A = Ea)$. Now we want $A^{-1} = A^{\odot}$.

(i)$\forall a \in A^{-1}$,for $A^{-1}A = E$,and $e \in E$,so $\exists b \in A^{-1}, b^{'} \in A$,s.t.$bb^{'} = e$,so $b^{1} = b^{-1} \in A$.And $A^{-1} = bE \Rightarrow (\exists c \in E)(a = bc) \Rightarrow (a^{-1} = c^{-1}b^{-1}eEA = A) \Rightarrow (a \in A^{\odot}$ *that is* $A^{-1} \subseteq A^{\odot})$.

(ii) $(\forall a \in A^{\odot}) \Rightarrow (a^{-1} \in A)$,and $(AA^{-1} = E)^{-1} \Rightarrow (\exists b \in A^{-1})a^{-1}b = e \Rightarrow a = b \in A^{-1}$,that is $a^{-1} \supseteq A^{\odot}$,so from all above, we get $A^{-1} = A^{\odot}$.

Corollary 3.1. *The classic quotient group G/N is an uniform power group.*

Proof. For $N \lhd G$, and N is the identity of G/N, so from theorem 3.4 the conclusion follows consequently.

Theorem 3.5. *If $\mathcal{G}$ is a power group of G and $e \in A \in \mathcal{G}$,then $A^{-1} \subset E \subset A$.*

Proof. For $e \in A$ and $A = AE = (\{e\} \cup A \backslash \{e\})E = E \cup (A \backslash \{e\})E$,so $E \subset A$,and $E = AA^{-1} = [E \cup (A \backslash E)]A^{-1} = A^{-1} \cup (A \backslash E)A^{-1}$,so $A^{-1} \subset E, A^{-1} \subset E \subset A$.

Theorem 3.6. *If G is a finite group, and $e \in A \in \mathcal{G}$,then $A^{-1} = E = A$.*

Proof. For G is finite, so A, E are both finite, from theorem 3.5 we know $A^{-1} \subset E \subset A$,and $|A| = |A^{-1}| \leq |E|$,thus $A^{-1} = E = A$.

Remark: it's very important to require G to be a finite group in this theorem, or else the conclusion does not hold. We can give an example that satisfy theorem 3.5 but does not comply with theorem 3.6: Let G be the real numbers addictive group, let $E = (0, +\infty)$ and H be the set of all integers, then $\mathcal{G} = \{n + E | n \in H\}$ is a power group of G, and E can be verified to be the identity, and $H \cong \mathcal{G}$,but $0 \in (-1, +\infty) = A \in \mathcal{G}$,and $0 \notin E$,so of course $A^{-1} = E = A$ does not hold. From the relationship of the elements of $\mathcal{G} : \cdots \supset (-2, +\infty) \supset (-1, +\infty) \supset E \supset (1, +\infty) \supset (2, +\infty) \supset \cdots$, we can tell that theorem 3.5 is satisfied.

Theorem 3.7. *If G is a finite group and $\mathcal{G}$ is a uniform power group of G,then $\forall A \in \mathcal{G}$,if $A \neq E$ indicate that $e \notin A$,then A forms a subgroup of G.*

Proof. For G is a finite group, and $e \in A$,so $A^{-1} = E = A$,and for $\mathcal{G}$ is a uniform power group, so E forms a subgroup of G , therefore A forms a subgroup of G.

Prove up. This is a short path to verify if A forms a subgroup of G.

Theorem 3.8. *Suppose G is a finite group, then the necessary condition for a subset $\mathcal{G}$ of the power set $\mathcal{P}(G)$ forms a power group is $\forall A \in \mathcal{G}, A \neq E \Rightarrow e \notin A$.*

The proof can be obtained directly from theorem 3.6.

This theorem claims that if $\mathcal{G}$ forms a power group of G , then there exists at most one element $E \in \mathcal{G}$ which contains the identity e of G. There are no other elements that contains e except for E .

Corollary 3.2. *Suppose G is a finite group, if $\mathcal{G} \subset \mathcal{P}(G)$ contains two sets A, B both of which contains e, then $\mathcal{G}$ is not able to form a power group of G.*

4 Characterization of Power Groups by Common Groups

Suppose G is a group,N is a normal group of G,G/N is a quotient group, and we use $\mathcal{G}|P(N)$ to denote the restriction of $\mathcal{G}$ on $P(N)$. In this section, we discuss G, N and the relationship between G/N and the corresponding power group, then we will get the characterization of power groups by common groups.

Theorem 4.1. *Suppose G is a group,N is a subgroup of G,$\mathcal{G}$ is one of the power groups of G, the identity is denoted by E , if $E \subset N$,then $\mathcal{G}|P(N)$ forms an power group of N so it's a subgroup of $\mathcal{G}$.*

Proof. (i) $\forall A, B \in \mathcal{G}|P(N)$,we have $A \bullet B \in \mathcal{G}$,and for $A \subset N, B \subset N \Rightarrow A \bullet B \subset N \subset A \bullet B \in P(N)$,thus $A \bullet B \in \mathcal{G}|P(N)$.

(ii) $\forall B \in \mathcal{G}|P(N)$,there exists $B^{-1} \in \mathcal{G}|P(N)$ s.t.$BB^{-1} = E$,so $\forall b^{'} \in B^{-1}, b \in B \subset N, \exists h \in E \subset N$,s.t.$b^{'}b = h \Rightarrow b^{'} = hb^{-1} \in N \Rightarrow B^{-1} \subset N \Rightarrow B^{-1} \in \mathcal{G}|P(N)$.

From (i),(ii) we know that $\mathcal{G}|P(N)$ is a power group of N , and it is consequently a subgroup of $\mathcal{G}$.

Theorem 4.2. *Suppose G is a group, N is a normal group of G,$\mathcal{G}$ is a power group of G,$E \subset N$,then*

(i) $\mathcal{G}|P(N)$ is a power group of N, so it's also a subgroup of $\mathcal{G}$.

(ii) $\mathcal{G}|P(N)$ is a normal subgroup of G.

Proof. (i) is a direct corollary of theorem 4.1.

(ii) $\forall A \in \mathcal{G}, \forall H \in \mathcal{G}|P(N)$, we only need to show that $AHA^{-1} \in \mathcal{G}|P(N)$. $\because AHA^{-1} \subset ANA^{-1} = NAA^{-1} = NE \subset N, \therefore AHA^{-1} \in P(N)$.And $\because AHA^{-1} \in \mathcal{G}, \therefore AHA^{-1} \in \mathcal{G}|P(N), \therefore \mathcal{G}|P(N)$ is a normal subgroup of $\mathcal{G}$.

Theorem 4.3. *Suppose G is a group, N is a normal group of G, then given any power group $\mathcal{G}$ on G, there is a certain subset of the power set of G/N.*

$$\varphi(G/N) = \{A^{'}|A^{'} \in P(G/N), A^{'} = \{aN|a \in A \in \mathcal{G}\}\},$$

and there exists an epimorphism f between $\mathcal{G}$ and $\varphi(G/N)$, thus $\varphi(G/N)$ is a group whose identity is $\{aN|a \in E\}$,the inverse of $\{aN|a \in A\}$ is $\{a|a \in A^{-1}\}$,and $\mathcal{G}/ker\ f \cong \varphi(G/N)$.

Proof. Let $f : \mathcal{G} \to \varphi(G/N), A \mapsto f(A) = \{aN|a \in A\}$,then $f(AB) = \{cN|c \in AB\}$, it's obvious that f is a surjection; for $f(A) \bullet f(B) = \{aN|a \in A\} \bullet \{bN|b \in B\} = \{abN|a \in A, b \in B\}$,and so $f(A)f(B) \subseteq f(AB)$.And for $\forall cN \in f(AB)$, we have $c \in AB \Rightarrow (\exists a \in A, b \in B)(c = ab) \Rightarrow (cN \in f(A)f(B)) \Rightarrow f(A) \bullet f(B) \supseteq f(AB), f(A)f(B) = f(AB)$,thus f is an epimorphism. Additionally, $\{aN|a \in E\}$ is the identity, and it's obvious that the inverse of $\{aN|a \in A\}$ is $\{aN|a \in A^{-1}\}$, it's unnecessary to go into details.

Theorem 4.4. *Suppose G is a group, N is its subgroup, if $E \subset N$ then*

$$\mathcal{G}/\mathcal{G}|P(N) \cong \varphi(G/N).$$

Proof. From theorem 4.3 we can get:$\mathcal{G}/ker\ f \cong \varphi(G/N)$, thus we just need to prove $\mathcal{G}|P(N) = ker\ f$.

(i) $\forall A = ker\ f, f(A) = \{aN|a \in A\} = \{cN|c \in E\} = f(E)$,so $\forall a \in A, \exists b \in E$,s.t. $aN = bN \Rightarrow (\exists n_1, n_2 \in N)(an_1 = bn_2) \Rightarrow a = bn_1n_2^{-1}$,so $b \in E \subset N, n_2n_1^{-1} \in N$,so $a \in N \Rightarrow A \in \mathcal{G}|P(N) \Rightarrow ker\ f \subseteq \mathcal{G}|P(N)$.(ii)$\forall A \in \mathcal{G}|P(N)$,then $A \subseteq N$,so $f(A) = \{aN|a \in A\} = \{aN|a \in EA\} = \{aN|a = ba_1 \in EA, b \in E, a_1 \in A\}$=$\{ba_1N|b \in E, a_1 \in A\} = \{bN|b \in E\} = f(E)$,so $\mathcal{G}|P(N) = ker\ f$,thus $\mathcal{G}|P(N) = ker\ f$, which means $\mathcal{G}/\mathcal{G}|P(N) \cong \varphi(G/N)$.

Notice that the condition that $E \subset N$ is quite important here, from this proof we can tell that $\mathcal{G}|P(N)$ forms a normal subgroup of $\mathcal{G}$, therefore (ii) of theorem 4.2 can be viewed as a corollary of this theorem.

Theorem 4.5. *If $\varphi(G/N) = \{A^{'}|A^{'} = \{aN|a \in A, A \in P(G)\}\}$ forms a power group of G/N,let $A^* = \bigcup\{aN/aN \in A^{'}\}, \mathcal{G}^* = \{A^*/A^{'} \in \varphi(G/N)\}$, then $\mathcal{G}^*$ is a power group of G, and $\varphi(G/N) \cong \mathcal{G}^*$.*

Proof. (i) Let $f^* : \varphi(G/N) \to \mathcal{G}, A^{'} \mapsto A^*$,then it's obvious that f is a surjection from $\varphi(G/N)$ to $\mathcal{G}^*$. Now we want to show that f is a homomorphism which means $f^*(A^{'} \bullet B^{'}) = f^*(A^{'})f^*(B^{'})$.

For $f^*(A^{'}) = A^*, f^*(B^{'}) = B^*$, so we just need to have $f^*(A^{'} \bullet B^{'}) = A^* \bullet B^*$ which is proved as follows:

(1) $\forall y \in f^*(A^{'}B^{'}) = \bigcup\{dN|dN \in A^{'} \bullet B^{'}\}, \exists doN = aN \bullet bN \in A^{'}B^{'}$,while $aN \in A^{'}, bN \in B^{'}$,s.t. $y \in doN = aN \bullet bN$,$\therefore \exists x \in aN, y \in bN$ s.t. $y = xy \in A^*B^*$,which means $f^*(A^{'} \bullet B^{'}) \subseteq A^* \bullet B^*$.(2) $\forall xy \in A^* \bullet B^*, \exists aN \in A^{'}, bN \in B^{'}$,s.t.$x \in aN, y \in bN$,$\therefore x \bullet y \in aN \bullet bN = abN \in A^{'} \bullet B^{'} \in \varphi(G/N) \Rightarrow xy \in f^*(A^{'} \bullet B^{'})$. From (1) (2) above we know $f^*(A^{'} \bullet B^{'}) = f^*(A^{'})f^*(B^{'})$.So $\varphi(G/N) \sim \mathcal{G}^*$, it certainly forms a power group of G.

(ii)Now we show that f^* is an injection.If $A^{'} \neq B^{'}$, then we may assume $\exists aN \in A^{'}$,but $aN \notin B^{'}$ and $\because a \in aN \Rightarrow a \in A^*$, from these we assert that $a \notin B^*$,otherwise $a \in bN \in B^{'} \Rightarrow aN = bN \in B^{'}$,contradiction occurred.$\therefore$ $A^* \neq B^*$,which means $\varphi(G/N) \cong \mathcal{G}^*$.

This theorem shows that some sort of $G/N's$ power group $\varphi(G/N)$ can be isomorphic to some certain type of $G's$ power group $\mathcal{G}$ and $\mathcal{G}$ can be related to some subgroups of G, thus this theorem is the one which joins the three aspects together. What's more, from some points of view, theorem 4.3 and theorem 4.6 have a similar sense to the fundamental theorem of isomorphism of classic algebra.

5 Structures of Power Groups

First, we consider the inverse proposition of 2.2: suppose $\mathcal{G}$ is a power group of G, does there exist a subgroup of G named H,s.t. $\mathcal{G} = \{aE|a \in H\}$?

Theorem 5.1. *Suppose $\mathcal{G}$ is a power group of G, if E is a subgroup of G,then $\mathcal{G} = \{aE|a \in G^*\}$ and G^* is a subgroup of G.*

Proof. $\forall A \in \mathcal{G}$,choose $a \in A$, then $aE \subset AE = A$.Now we prove $aE = A$.If not, there must be $b \in A \backslash aE$, we can show that $a^{-1}b \notin E$.For if $a^{-1}b = c \in E$,then $b = ac \in aE$,this leads to a contradiction for it's mentioned before that $b \in A \backslash aE$.Let $d \in A^{-1}$,then $da, db \in A^{-1}A = E$,so $a^{-1}b = a^{-1}(d^{-1}d)b(da)^{-1}(db) \in E$(for E is a subgroup), this goes against the fact that $a^{-1}b \notin E$.Thus $aE = A$,that is $\mathcal{G} \subset \{aE|a \in G^*\}$.Otherwise,$\forall a \in G^*, \exists A \in \mathcal{G}$,s.t. $a \in A$,so $aE = A \in \mathcal{G}, \{aE|a \in G^*\} \subset \mathcal{G}$.Then we will prove that G^* is a subgroup of G.

$\forall a \in G^*, \exists a \in \mathcal{G}$,let $a \in A$, notice that $e \in E$ and $AA^{-1} = E$,so $\exists b \in A, b^{-1} \in A^{-1}$ s.t. $bb^{-1} = e$. Then from $A = bE$ we know that $\exists c \in E$,s.t. $a = bc$,so $a^{-1} = (bc)^{-1} = c^{-1}b^{-1} \in EA^{-1} = A^{-1} \subset G^*$.And from theorem 3.1 1) we know that $G^* \leq G$,$\forall x \in A^{-1}$,$A^{-1} = xE = a^{-1}(ax)E, ax \in AA^{-1} = E$,so $A^{-1} = a^{-1}E \Rightarrow a^{-1} \in A^{-1}$,that is $A^{-1} \in G^*$.

Under the condition of this theorem, we have three corollaries listed below

1)E is a normal subgroup of G^*;2) $G^* = G^0$;3)$\mathcal{G} = G^*/E$.

In addition, there are three special cases listed below:

1)G is a periodic group $\Rightarrow \mathcal{G} = G^*/E$;2) E is finite $\Rightarrow \mathcal{G} = G^*/E$;3)$G$ is finite $\Rightarrow \mathcal{G} = G^*/E$.

Theorem 5.1 answered the problem raised at the beginning of this section affirmatively under the condition that E is a subgroup, but this condition is so strong that $\mathcal{G}$ is strengthened to a quotient group G^*/E. Now we will loosen the constraint conditions and mainly consider the construction of regular groups.

Definition 5.1. *Set S to be a monoid of G (the subsemigroup which contains the identity of G),H is a subgroup of G, we call S the normal subsemigroup of G with respect to H, if $(\forall a \in H)(aS = Sa)$. Especially, when $H = G$, we can say that S is a normal group of G.*

If $\mathcal{G}$ is a regular power group of G, we can guess that E is a regular subsemigroup of some kind of subgroup of G.

Suppose $\mathcal{G}$ is a power group of G, $\forall A \in \mathcal{G}$,let $\overline{A} \triangleq \{a \in A, a^{-1} \in A^{-1}\}$ which is called the kernel of A. Set $\overline{G} \triangleq \bigcup\{\overline{A}|A \in \mathcal{G}\}$, it's easy to verify this two properties:1)$e \in E \Rightarrow (\forall A \in \mathcal{G})(\overline{A} \neq \varnothing)$;Otherwise $(\exists A \in \mathcal{G})(\overline{A} \neq \varnothing) \Rightarrow e \in E$; 2) $\overline{G} \neq \varnothing \Leftrightarrow e \in E$.

Theorem 5.2. *If $\mathcal{G}$ is the regular power group of G, then $\overline{G}$ is a subgroup of G^0 and $\mathcal{G} = \{aE|a \in \overline{G}\}$.*

Proof. $\forall a, b \in \overline{G}, \exists A, B \in \mathcal{G}$, choose $a \in \overline{A}, b \in \overline{B}$,then $ab^{-1} \in AB^{-1} = C \in \mathcal{G}$. Notice that $(ab^{-1})^{-1} = ba^{-1} \in BA^{-1} = C^{-1}$,so $ab^{-1} \in \overline{C} \subset \overline{G}$ which means $\overline{G}$ is a subgroup of G^0. What else, $\forall A \in \mathcal{G}$,for $\mathcal{G}$ is regular, so $A \neq \varnothing$,choose $a \in \overline{A}$,we have $aE = A$. In fact, $b \in A \mapsto b = eb = (aa^{-1})b = a(a^{-1}b) \in a(A^{-1}A) = aE$,so $A \subset aE$;for $aE \subset AE = A$,so $A = aE$,thus $\mathcal{G} \subset \{aE|a \in \overline{G}\}$ .Otherwise,$\forall a \in \overline{G}, \exists A \in \mathcal{G}$,let $A = aE$, that is $\{aE|a \in \overline{G}\} \subset \mathcal{G}$.

Under the conditions of theorem 5.2, we have direct corollaries$E$ is a normal subsemigroup of $\overline{G}$ and $\overline{E}$ is a normal subgroup of $\overline{G}$.

Theorem 5.2 answered the question raised in this section under the condition that $\mathcal{G}$ is regular, next we will describe the construction of power group in more details.

Definition 5.2. *If E is a normal subsemigroup of G with respect to H, denote $H|E \triangleq \{aE|a \in H\}$, then $H|E$ is a regular power group of G, called a quasi quotient group of H with respect of E. Especially, when $H = G$, call $G|E$ a quotient group of G(with respect to E).*

According to this definition, theorem 5.2 means that if $\mathcal{G}$ is regular, the $\mathcal{G}$ must be some quasi quotient group of G with respect to E: $\mathcal{G} = \overline{G}|E$. We may say that $\overline{G}|E$ and $\overline{G}/\overline{E}$ are isomorphic.

Theorem 5.3. *If $\mathcal{G}$ is a regular power group of G, then $\overline{G}/\overline{E} \cong \overline{G}|E$.*

Proof. Let $f : \overline{G} \to \overline{G}/E, a \mapsto aE$. It's easy to see that f is an epimorphism, thus $\overline{G}/ker\ f \cong \overline{G}|E$, now we want to prove that $ker\ f = \overline{E}$.$\because a \in ker\ f \Rightarrow a = ae \in aE = E \Rightarrow (\exists b \in E)(ab = e) \Rightarrow a^{-1} = b \in E \Rightarrow a \in E$, in addition, notice that the proof of theorem 5.2 shows the fact that $(\forall A \in \mathcal{G})(a \in \overline{A} \Rightarrow aE = A = Ea)$, thus $a \in \overline{E} \Rightarrow aE = E \Rightarrow a \in ker\ f$.

This theorem makes it clear what the structure of power group $\mathcal{G}$ is like: $\mathcal{G} = \overline{G}|E \cong \overline{G}/\overline{E}$. Let us study the structure of $\overline{G}/\overline{E}$.Suppose $\mathcal{G}$ is a power group of G, and $\mathcal{G} \triangleq \{\overline{A}|A \in \mathcal{G}\}$.

Theorem 5.4. *If $\mathcal{G}$ is a regular power group of G, then $\mathcal{G} = \overline{G}/\overline{E}$.*

Proof. $\overline{A} \in \overline{\mathcal{G}} \Rightarrow (\forall a \in \overline{A})(A = aE) \Rightarrow (\forall a \in \overline{A})(\overline{A} = \overline{aE})$, next we want to show $\overline{aE} \cong a\overline{E}$.For $x \in \overline{aE} \Rightarrow (\exists h \in E)(x = ah) \Rightarrow h^{-1} = x^{-1}a \in (aE)^{-1}A = A^{-1}A = E \Rightarrow h \in \overline{E} \Rightarrow x = ah \in a\overline{E}$.Otherwise,$x \in a\overline{E} \Rightarrow (\exists h \in \overline{E})(x = ah) \Rightarrow x^{-1} = h^{-1}a^{-1} \in Ea^{-1} = a^{-1}E = (aE)^{-1} \Rightarrow x \in \overline{aE}$. In a word, $\overline{aE} = a\overline{E}$,this conclusion indicates that $\overline{\mathcal{G}} \subset \overline{G}/\overline{E}$.On another hand, $\forall a\overline{E} \in \overline{G}/\overline{E}$,for $a \in \overline{G}$,then $\exists \overline{A} \in \overline{\mathcal{G}}$,s.t.$a \in \overline{A}$,thus $\overline{A} = \overline{aE} = a\overline{E}$, which means $\overline{G}/\overline{E} \subset \overline{\mathcal{G}}$.

6 Conclusion

Since Prof. Wang Peizhuang raised the problem of topological upgrading and obtained a series of systematic theoretical results and applications, and hypergroup which was raised by H.S.Wall had a great number of important applications in theoretical physics including quantum mechanics and string theory, this kind of upgrading of mathematical structure had become especially important. This paper studies this problem, namely the upgrading from group structure to its power set. In specific, we give the definition, examples and ways of construction of power groups, then show the relationship between power groups and common groups and characterized power groups by common groups and obtain a series of constructive conclusions. So for, the breakthroughs of upgrading a group to its power set have been achieved.

Acknowledgements. This research is supported by the National Natural Science Foundation of China (No. 70771030) and by the Natural Science Foundation of Guangdong Province (No. 8451009101001040).

References

1. Wang, P.-Z.: The falling-shadow of fuzzy set and random set. Beijing Normal University Press, Beijing (1984)
2. Pei-zhuang Wang, E. Sanchez: Treating a fuzzy subset as a projectable random subset. Fuzzy Information and Decision 2, 213–219 (1982)
3. Wang, P.-z., Sugeno, M.: The factor fields and background structure for fuzzy sets. Fuzzy Math. 2, 45–54 (1982)
4. Wang, P.-z.: From the fuzzy statistics to the falling random subsets. Advance in Fuzzy set Theory and Appl. 1, 81–96 (1983)
5. Wang, P.-z.: Random sets and fuzzy sets. In: International Encyclopedia on Systems and Control. Pergamon Press, Oxford (1983)
6. Wang, P.-z.: Fuzzy contactability and fuzzy varlables. Fuzzy Sets and Systems 8(1), 81–92 (1982)
7. Wang, P.-z., Sanchez, E.: Hyperfields and random sets. In: Proc. of IFAC symposium on Fuzzy Information, Knowledge Representation and Decision Analysis, pp. 325–330 (1983)

8. Wang, P.-z., Liu, X.-h., Chen, Y.-y.: Fuzzy force and fuzzy dynamic analysis. In: Proc. of IEEE Symposium on Man, Machine and Cyber. (1984)
9. Wang, P.-z., Liu, X.-h., Chuan, K.: Set-valued Statistics. In: Proc. of FIP 1984 Symposium on fuzzy Information Processing (1984)
10. Wall, H.S.: Hypergroups. American Journal of Mathematics 59(1), 77–98 (1937)
11. Rosenberg, I.G.: An Algebraic Approach to Hyperalgebras. IEEE, Los Alamitos (1996)
12. Robinson, D.J.S.: A Course in the Theory of Groups. Springer, New York (1995)
13. Li, H.-x., Wang, P.-z.: The Power Group. Applied Mathematics 4(1), 1–4 (1988)
14. Zhong, Y.-b.: The Structure and Relationship on Hypergroup. Chinese Quarterly Journal of Mathematics 5(4), 102–106 (1990)
15. Zhong, Y.-b.: The Structure of Hypergroup on the Cyclical Group. Chinese Quarterly Journal of Mathematics 9(1), 26–31 (1994)
16. Zhong, Y.-b.: Isomorphism on Hypergroup. Chinese Quarterly Journal of Mathematics 9(2), 60–63 (1994)
17. Zhong, Y.-b.: The Homomorphism and Isomorphism on Hypergroup. Chinese Quarterly Journal of Mathematics 12(3), 7–14 (1997)
18. Zhong, Y.-b.: The cardinal number theorem of fuzzy hypergroup. Fuzzy System and Mathematics 12(2), 61–69 (1998)
19. Zhong, Y.-b.: The Existence of HX-RING. Applied Mathematics Journal of Chinese University 15(2), 134–138 (2000)
20. Zhong, Y.-b.: Fuzzy Hyperring. Journal of Guangzhou University 5(4), 1–5 (2006)
21. Zhong, Y.-b.: Fuzzy Hyper-Topological Group. Advance in soft computing 54(1), 324–330 (2009)

P-Sets and Applications of Internal-Outer Data Circle

Li Zhang[1,2], Ming Xiu[2], and Kai-quan Shi[1]

[1] Department of Mathematics and System Sciences,
Shandong University, Jinan 250100, China
e-mail: xitonglilun@126.com
[2] College of Engineering Technology, Jimei University,
Xiamen 361021, China
e-mail: jdxm1@tom.com

Abstract. P-sets (packet sets) which come from finite general sets have new mathematic structure. Introduced to dynamic characteristic, P-sets are composed of internal P-sets $X^{\bar{F}}$ (internal packet set $X^{\bar{F}}$) and outer P-sets X^{F} (outer packet sets X^{F}). P-sets have dynamic characteristics. Based on P-sets, the concepts of internal-data circle, outer-data circle and data circle theorem are given. By these results, the dynamic data restore theorem, dynamic data restore guideline and dynamic data restore-identification are given. Applications in dynamic data analysis-identification are proposed. P-sets can be used in many fields.

Keywords: P-sets, dynamic characteristic, internal-outer data circle, data restore, data restore theorem, data circle theorem, applications.

1 Introduction

Data transmission system of computer (Perspective identification system of computer) often appears the phenomenon as this: the data which system normally outputs is $m=\{x_1,x_2,\ \cdots,x_q\}$ .If internal parameters of information transmission module network change (module device is aging, failure), data m changes into $m^{\bar{F}}=\{x_1,x_2,\cdots,x_p\}, p<q,$ or data m changes into $m^{F}=\{x_1,x_2,\cdots,x_r\}, q<r,$ the performance of system state is that : the former, some data in m has been lost (m has lost many data elements x_i), the latter, some data has been added to m (m has been supplemented by some data elements x_j), the output given by system is "a chaotic image distortion". The abnormal state of system is similar to the characteristic of P-sets. This phenomenon has been commonplace, the theoretical understanding of this phenomenon is given, we can not find the results of such research in the existing international and domestic literature, how many results the phenomenon has hidden in the end?

Refs. [1,2] proposed P-sets and the structure of P-sets was given. P-sets are a set pair which are composed of internal P-sets $X^{\bar{F}}$ (internal packet sets $X^{\bar{F}}$) and outer P-sets X^{F} (outer packet sets X^{F}), or $(X^{\bar{F}},X^{F})$ are P-sets. P-sets have dynamic

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 581–591.
springerlink.com

characteristic [1-9]. Based on the structure of P-sets and dynamic characteristic, the paper gives discussion. Having a theoretical understanding of the nature about this phenomenon is the subject of this paper.

By using P-sets, the concepts of internal data-circle and outer data-circle are given, Based on these concepts, data restore guideline and data restore-identification theorem are proposed. Finally application is given. It is important to understand the “pathological” nature of data transmission network. P-sets are new theory and method to study the dynamic information system.

For the convenience of discussion and accepting the results of this paper easily while keeping the contents integral, the structure of P-sets is simply introduced to section 1 as the theory basis of the discussion of this paper. The more concepts and applications of P-sets can be found in Refs. [1-9].

2 P-Sets and Its Set Pair Structure

Assumption 1. *X is a finite general set on U, U is a finite element universe, V is a finite attribute universe.*

Given a general set $X=\{x_1,x_2,\cdots,x_m\}\subset U$, and $\alpha=\{\alpha_1,\alpha_2,\cdots,\alpha_k\}\subset V$ is attribute set of X, $X^{\overline{F}}$ is called internal packet sets of X, called internal P-sets for short, moreover

$$X^{\overline{F}}=X-X^{-}, \tag{1}$$

X^{-} is called $\overline{F}$-element deleted set of X, moreover

$$X^{-}=\{x\mid x\in X,\overline{f}(x)=u\,\overline{\in}\,X,\overline{f}\in\overline{F}\}, \tag{2}$$

if the attribute set α^{F} of $X^{\overline{F}}$ satisfies

$$\alpha^{F}=\alpha\cup\{\alpha'\mid f(\beta)=\alpha'\in\alpha,f\in F\}, \tag{3}$$

where $X^{\overline{F}}\neq\phi,\beta\in V,\beta\,\overline{\in}\,\alpha,f\in F$ turns β into $f(\beta)=\alpha'\in\alpha$.

Given a general set $X=\{x_1,x_2,\cdots,x_m\}\subset U$, and $\alpha=\{\alpha_1,\alpha_2,\cdots,\alpha_k\}\subset V$ is attribute set of X, X^{F} is called outer packet sets of X, called outer P-sets for short, moreover

$$X^{F}=X\cup X^{+}, \tag{4}$$

X^{+} is called F-element supplemented set, moreover

$$X^{+}=\{u\mid u\in U,u\,\overline{\in}\,X,f(u)\in x'\in X,f\in F\}, \tag{5}$$

if the attribute set $\alpha^{\overline{F}}$ of X^{F} satisfies

$$\alpha^{\overline{F}}=\alpha-\{\beta_i\mid\overline{f}(\alpha_i)=\beta_i\,\overline{\in}\,\alpha,\overline{f}\in\overline{F}\}, \tag{6}$$

where $\alpha^{\overline{F}}\neq\phi,\alpha_i\in\alpha,\overline{f}\in\overline{F}$ turns α_i into $\overline{f}(\alpha_i)=\beta_i\,\overline{\in}\,\alpha$.

The set pair which are composed of internal P-sets $X^{\bar{F}}$ and outer P-sets X^{F} are called P-sets (packet sets) generated by general set X, called P-sets for short, moreover

$$(X^{\bar{F}}, X^{F}), \tag{7}$$

where general set X is ground set of $(X^{\bar{F}}, X^{F})$.

As P-sets have dynamic characteristic, the general representation of P-sets is:

$$\{(X_i^{\bar{F}}, X_j^{F}) \mid i \in \mathrm{I}, j \in \mathrm{J}\}, \tag{8}$$

where I, J are index sets, formula (8) is the representation of set pair family of P-sets. In Fig. 1, $X \subset U$ is a finite general set on $X = \{x_1, x_2, \cdots, x_q\}$, $X^{\bar{F}}$ is internal P-sets of $X \subset U, X^{\bar{F}} = \{x_1, x_2, \cdots, x_p\}, p \le q, X^{F}$ is outer P-sets of $X \subset U, X^{F} = \{x_1, x_2, \cdots, x_r\}, q \le r$, $(X^{\bar{F}}, X^{F})$ is P-sets. $X^{\bar{F}}$ is expressed in thick line, X is expressed in thin line, X^{F} is expressed in dashed line.

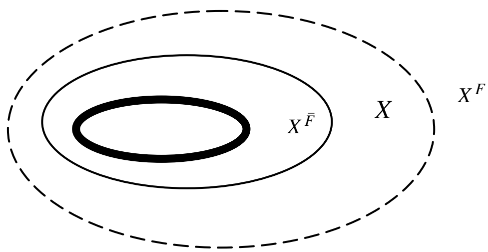


Fig. 1. Shows intuitive graphical representation of P-sets

Important Description on the structure and concepts of P-sets

1°. For the convenience and without misunderstanding, formula (7) only uses one of many set pairs to express P-sets.

2°. $F = \{f_1, f_2, \cdots, f_m\}$, $\bar{F} = \{\bar{f}_1, \bar{f}_2, \cdots, \bar{f}_n\}$ are element transfer families, $f \in F, \bar{f} \in \bar{F}$ are element transfers, $f \in F, \bar{f} \in \bar{F}$ are given functions (function is a transformation or mapping). The characteristic of $\bar{f} \in \bar{F}$ is that $u \in U, u \,\bar{\in}\, X, f \in F$ changes u into $f(u) = x' \in X$, or $\beta \in V, \beta \,\bar{\in}\, \alpha, f \in F$ changes β into $f(\beta) = \alpha' \in \alpha$, the characteristic of $f \in F$ is that $x \in X$, $\bar{f} \in \bar{F}$ changes x into $\bar{f}(x) = u \,\bar{\in}\, X$, or $\alpha_i \in \alpha$, $\bar{f} \in \bar{F}$ changes α_i into $\bar{f}(\alpha_i) = \beta_i \,\bar{\in}\, \alpha$.

3°. The characteristic of formula (3) is similar to the structure of $T = T + 1$, $T = T + 1$ has dynamic characteristic, it is the same with formula (3), $\{\alpha' \mid f(\beta) = \alpha' \in \alpha, f \in F\}$ means the set composed of new elements added to α, $\{\alpha' \mid f(\beta) = \alpha' \in \alpha, f \in F\}$ and α which has not been supplemented by elements satisfy that $\{\alpha' \mid f(\beta) = \alpha' \in \alpha, f \in F\} \cap \alpha = \phi$.

4°. The dynamic characteristic of formula (4) is that formulae (4) and (5) can be denoted by a common equation as follow:

$$X^F = X \bigcup \{u \mid u \in U, u \overline{\in} X, f(u) = x' \in X, f \in F\}\,. \tag{9}$$

There is $u_1 \in U, u_1 \overline{\in} X, f(u_1) = x_1' \in X$, formula (9) changes into $X_1^F = X \bigcup \{u \mid u \in U, u \overline{\in} X, f(u) = x' \in X, f \in F\} = X \bigcup \{x_1'\} = \{X, x_1'\}$. Suppose that $X = X_1^F$, taking $u_2, u_3 \in U, u_2, u_3 \overline{\in} X, f(u_2) = x_2' \in X, f(u_3) = x_3' \in X$, formula (9) changes into $X_2^F = X_1^F \bigcup \{u \mid u \in U, u \overline{\in} X, f(u) = x' \in X, f \in F\} = X \bigcup \{x_2', x_3'\} = \{X, x_1'\} \bigcup \{x_2', x_3'\} = \{X, x_1', x_2', x_3'\}$ .Suppose that $X = X_2^F$, taking $u_4 \in U; u_4 \overline{\in} X;\ f(u_4) = x_4' \in X$, formula (9) changes into $X_3^F = X_2^F \bigcup \{u \mid u \in U, u \overline{\in} X, f(u) = x' \in X, f \in F\} = X_2^F \bigcup \{x_4'\} = X_1^F \bigcup \{x_2', x_3'\} \bigcup \{x_4'\} = X \bigcup \{x_1'\} \bigcup \{x_2', x_3'\} \bigcup \{x_4'\} = \{X, x_1', x_2', x_3', x_4'\}$, and so on , then there is $X_1^F \subset X_2^F \subset X_3^F$, or, card$(X_1^F) \le$ card$(X_2^F) \le$ card(X_3^F), card=cardinal number. Outer P-sets X^F are larger. Formula (9) is similar to the characteristic of $T = T + 1$.

5°. In formulae (1)-(3), some elements are deleted from X , X generates internal P-sets $X^{\overline{F}}$, it is the same with that the attribute set α of X is supplemented by some elements, α generates $\alpha^F, \alpha \subset \alpha^F$ .Or, if α_1^F, α_2^F are attribute sets of $X_1^{\overline{F}}$, $X_2^{\overline{F}}$ respectively, or $\alpha_1^F \subseteq \alpha_2^F$,then $X_2^{\overline{F}} \subseteq X_1^{\overline{F}}$. $\{\alpha' \mid f(\beta) = \alpha' \in \alpha, f \in F\}$ in formula (3) is not attribute set of X^- which is composed of elements deleted from X , the denotation of X^- is formula (2).

Based on formulae (1)-(7), there is relation between P-sets $(X^{\overline{F}}, X^F)$ and general set X .

Theorem 1. *If* $\overline{F} = F = \phi$ *, then P-sets* $(X^{\overline{F}}, X^F)$ *and general set* X *satisfy*

$$(X^{\overline{F}}, X^F)_{\overline{F}=F=\varnothing} = X\ . \tag{10}$$

Proof. If $\overline{F} = \phi$, then formula (2) changes into $X^- = \{x \mid x \in X, \overline{f}(x) = u \overline{\in} X, \overline{f} \in \overline{F}\} = \phi$, formula (1) changes into $X^{\overline{F}} = X - X^- = X$, If $F = \phi$, then formula (5) changes into $X^+ = \{u \mid u \in U, u \overline{\in} X, f(u) = x' \in X, f \in F\} = \phi$, formula (4) changes into $X^F = X \bigcup X^+ = X$. P-sets $(X^{\overline{F}}, X^F)$ change into X , then there is formula (10).

Formula (10) proposes that under the condition of $\overline{F} = F = \phi$, P-sets $(X^{\overline{F}}, X^F)$ turn back to "origin" of general set , in other words, P-sets have lost dynamic characteristics, actually, P-sets $(X^{\overline{F}}, X^F)$ are general sets X .

Theorem 2. *If* $\overline{F} = F = \phi$ *, the set pair family of P-sets* $\{(X_i^{\overline{F}}, X_j^F) \mid i \in \mathrm{I}, j \in \mathrm{J}\}$ *and general set* X *satisfy:*

$$\{(X_i^{\overline{F}}, X_j^F) \mid i \in \mathrm{I}, j \in \mathrm{J}\}_{\overline{F}=F=\varnothing} = X\ . \tag{11}$$

Formula (11) proposes that under the condition of $\bar{F} = F = \phi$, each $X_i^{\bar{F}}$, X_j^F turns back to "origin" of general set , or $\{(X_i^{\bar{F}}, X_j^F) \mid i \in \mathrm{I}, j \in \mathrm{J}\}$ turns back to "origin" of general set . The proof of theorem 2 is similar to theorem 1, so the proof is omitted.

3 Internal-Outer Data Circle and Dynamic Data Restore

Assumption 2. $X, X^{\bar{F}}$, and X^F, in section 1 are denoted by $m, m^{\bar{F}}$, and m^F ,respectively, or, $m = X, m^{\bar{F}} = X^{\bar{F}}$, and $m^F = X^F$, in order to avoid confusion and misunderstanding.

Definition 1. *$m^{\bar{F}}$ is called $\bar{F}$-data generated by data m ,if*

$$m^{\bar{F}} = \{x_1, x_2, \cdots, x_p\}, \tag{12}$$

$\forall x_i \in m^{\bar{F}}$ is called data element of $m^{\bar{F}}$.

Definition 2. *m^F is called F-data generated by data m ,if*

$$m^F = \{x_1, x_2, \cdots, x_r\}, \tag{13}$$

$\forall x_j \in m^F$ is called data element of m^F,
where $m = \{x_1, x_2, \cdots, x_q\}$, p, r in formulae (11), (12) satisfy that $p \le q \le r$. $p, q, r \in \mathrm{N}^+$.

Definition 3. *$y^{\bar{F}}$ is called characteristic value set of $m^{\bar{F}}$, if*

$$y^{\bar{F}} = \{y_1, y_2, \cdots, y_p\}, \tag{14}$$

y^F is called characteristic value set of m^F, if

$$y^F = \{y_1, y_2, \cdots y_r\} , \tag{15}$$

where $\forall y_k \in y^{\bar{F}}$ is characteristic value of data element $x_k \in m^{\bar{F}}$ (the value of x_k, or, the value of x_k that system outputs), $\forall y_\lambda \in y^F$ is characteristic value of data element $x_\lambda \in m^F$, $y_k, y_\lambda \in \mathrm{R}$, R is real number set.

Definition 4. O_γ *is called data unit circle which considers coordinate origin O as the center and considers ρ as the radius ,called data circle for short, if*

$$\rho = \| y \| / \| y \|, \tag{16}$$

where $\| y \| = (y_1^2 + y_2^2 + \cdots + y_q^2)^{1/2}$ is 2-Norm of vector $y = (y_1, y_2, \cdots, y_q)^T$ generated by characteristic value set $y = \{y_1, y_2, \cdots, y_q\}, y = \{y_1, y_2, \cdots, y_q\}$ is characteristic value set of data $m = \{x_1, x_2, \cdots, x_q\}$.

Definition 5. $O^{\bar{F}}$ *is called* $\bar{F}$ *-data circle which considers coordinate origin O as the center and considers* $\rho^{\bar{F}}$ *as the radius, if*

$$\rho^{\bar{F}} = \| y^{\bar{F}} \| / \| y \| , \tag{17}$$

where $\| y^{\bar{F}} \| = (y_1^2 + y_2^2 + \cdots + y_p^2)^{1/2}$ *is 2-Norm of vector* $y^{\bar{F}} = (y_1, y_2, \cdots, y_p)^T$ *,* $y^{\bar{F}}$ *is vector generated by formula (14).*

Definition 6. O^F *is called* F *-data circle which considers coordinate origin O as the center and considers* ρ^F *as the radius , if*

$$\rho^F = \| y^F \| / \| y \| , \tag{18}$$

where $\| y^F \| = (y_1^2 + y_2^2 + \cdots + y_r^2)^{1/2}$ *is 2-Norm of vector* $y^F = (y_1, y_2, \cdots, y_r)^T$ *,* y^F *is vector generated by formula (15).*

Obviously, $m^{\bar{F}}$ in formula (12) comes from data m which has lost some data elements x_i , m^F in formula (13) comes from data m which has been supplemented by some data elements x_j , formulae (12) and (13) are two states of phenomenon given by the introduction, and they are the "sick out" of system. By definitions 1-6, we can get that:

Theorem 3 (Data circle theorem of steady data). *If the output* m_k, m_λ *are the data which system* w *outputs when at time* k, $\lambda \in T$, $k \neq \lambda$ *,and they satisfy*

$$\text{UNI}\{m_k, m_\lambda\} , \tag{19}$$

then the data circle O_k *generated by* m_k *and* O_λ *generated by* m_λ *coincide ,or*

$$O_k = O_\lambda , \tag{20}$$

where UNI=unidentification[1,2]*.*

We can get the proof easily, so the proof of theorem 3 is omitted, or it can be obtained from Fig. 2 straightforwardly.

Corollary 1. *If the output of system* w *is steady, then the output* m_t *of system* w *at any time* t *constitutes data circle* O .

Theorem 4 (Data circle internal-concentric circle theorem). *If the output of system* w *is* $\bar{F}$ *-data* $m^{\bar{F}}$ *, the data circle* $O^{\bar{F}}$ *generated by* $m^{\bar{F}}$ *is internal-concentric circle of* O *, or*

$$O^{\bar{F}} \subset O , \tag{21}$$

where" $\subset$ *"denotes that* $O^{\bar{F}}$ *is in data circle* O .

In fact, from formulae (14), (17) and (16), we can get that $m^{\bar{F}}=\{x_1,x_2,\cdots,x_p\}\subseteq\{x_1,x_2,\cdots,x_q\}=m,\rho^{\bar{F}}=\|y^{\bar{F}}\|/\|y\|\le\|y\|/\|y\|=\rho$, or $0<\rho^{\bar{F}}\le\rho=1$, so there is formula (21).

Theorem 5 (Data circle outer-concentric circle theorem). *If the output of system w is F-data m^F, the data circle O^F generated by m^F is outer-concentric circle of data circle O, or*

$$\mathrm{O}\subset\mathrm{O}^F, \tag{22}$$

where "$\subset$" denotes that O^F is out of data circle O.

Using definitions 1-6, theorems 3-5 and corollary 1,we can get that :

Proposition 1. *The data of system w outputs at time $k\in T$ constitutes $\bar{F}$-data circle $\mathrm{O}^{\bar{F}}$, at time $k\in T$ system w has lost data, and vice versa.*

Proposition 2. *The data of system w outputs at time $\lambda\in T$ constitutes F-data circle O^F, at time $\lambda\in T$ the output of system w occurs data intrusion (Interfere data comes into system), and vice versa.*

Theorem 6 ($\bar{F}$-data restore theorem). *The necessary and sufficient condition of the data $m^{\bar{F}}$ which system w outputs being restored into m is that the attribute set α^F of data $m^{\bar{F}}$ and the attribute set α of data m satisfy*

$$(\alpha^F-\{\alpha_i\mid\alpha_i\in\alpha^F,\bar{f}(\alpha_i)=\beta_i\,\bar{\in}\,\alpha^F\})-\alpha=\phi\ . \tag{23}$$

Proof. In important description 5° of section 1: If $X_j^{\bar{F}}\subseteq X_i^{\bar{F}}$, the attribute set α_i^F of $X_i^{\bar{F}}$ and the attribute set α_j^F of $X_j^{\bar{F}}$ satisfy that $\alpha_i^F\subseteq\alpha_j^F$; or, if $\mathrm{card}(X_j^{\bar{F}})\le\mathrm{card}(X_i^{\bar{F}})$, there is $\mathrm{card}(\alpha_i^{\bar{F}})\le\mathrm{card}(\alpha_j^{\bar{F}})$,card=cardinal number. 1°. As $m^{\bar{F}}$ is $\bar{F}$-data of m, $m^{\bar{F}}\subseteq m$,the attribute set α^F of $m^{\bar{F}}$ and the attribute set α of m satisfy that $\alpha\subseteq\alpha^F$. If $m^{\bar{F}}$ is restored into m, or $m^{\bar{F}}=m$, $m^{\bar{F}}$ has the same attribute set with m. Obviously, there is attribute difference set $\nabla\alpha^F=\{\alpha_i\mid\alpha_i\in\alpha^F,\bar{f}(\alpha_i)=\beta_i\,\bar{\in}\,\alpha^F\}$, $\nabla\alpha^F$ being deleted from α^F makes there is formula $(\alpha^F-\{\alpha_i\mid\alpha_i\in\alpha^F,\bar{f}(\alpha_i)=\beta_i\,\bar{\in}\,\alpha^F\})-\alpha=\phi$, then $m^{\bar{F}}=m$. 2°. If there is $(\alpha^F-\{\alpha_i\mid\alpha_i\in\alpha^F,\bar{f}(\alpha_i)=\beta_i\,\bar{\in}\,\alpha^F\})-\alpha=(\alpha^F-\nabla\alpha^F)-\alpha=\phi$, or, $\alpha^F-\nabla\alpha^F=\alpha$, there must be $m^{\bar{F}}=m, m^{\bar{F}}$ is restored into m.

Theorem 7 (F-data restore theorem). *The necessary and sufficient condition of the data m^F which system w outputs being restored into m is that the attribute set $\alpha^{\bar{F}}$ of data m^F and the attribute set α of data m satisfy*

$$(\alpha^{\bar{F}} \cup \{\beta_i \mid \beta_i \in V, \beta_i \bar{\in} \alpha^{\bar{F}}, \bar{f}(\beta_i) = \alpha_i' \in \alpha^{\bar{F}}\}) - \alpha = \phi . \tag{24}$$

The proof is similar to theorem 6, so it is omitted.

Theorem 8 ($\bar{F}$ -Data restore-identification-Theorem). *If $\bar{F}$ -data circle $\bigcirc^{\bar{F}}$ and data circle $\bigcirc$ of m satisfy*

$$\bigcirc^{\bar{F}} = \bigcirc , \tag{25}$$

then data $m^{\bar{F}}$ which constitutes $\bar{F}$ -data circle $\bigcirc^{\bar{F}}$ is restored into data m which constitutes data circle $\bigcirc$, moreover

$$\text{UNI}\{m^{\bar{F}}, m\} . \tag{26}$$

Theorem 9 (F -Data restore-identification-Theorem). *If F -data circle $\bigcirc^{F}$ and data circle $\bigcirc$ of m satisfy*

$$\bigcirc^{F} = \bigcirc , \tag{27}$$

then data m^{F} which constitutes F -data circle $\bigcirc^{F}$ is restored into data m which constitutes data circle $\bigcirc$, moreover

$$\text{UNI}\{m^{F}, m\} . \tag{28}$$

Theorems 8, 9, can be obtained from theorems 6, 7 directly, so the proof is omitted.

$\bar{F}$ -data and F -data Restore guideline

If the output data m of system w at time $t \in T$ generates a data unit circle $\bigcirc$, the difference data $\nabla m^{\bar{F}}$, Δm^{F} satisfy

$$\nabla m^{\bar{F}} - \Delta m^{F} = \phi , \tag{29}$$

where $\nabla m^{\bar{F}} = m - m^{\bar{F}}, \Delta m^{F} = m^{F} - m$.

The guideline points out that the output data m of system w at time $t \in T$ does not exist data losing and data invasion.

By using the results in section 2, section 3 gives application of internal-outer data circle in data identification about dynamic output of system.

4 Data Identification of System State and Its Application

Example of this section is from computer vision identification system, data of the example is from raw data which is processed through the technical means, it does not affect the analysis of results .For simple and without losing generality, the example only gives $\bar{F}$ - data and application of $\bar{F}$ - data circle in data identification system state, the block diagram of system is omitted. Vision identification system w of computer has seven data-output terminals, its output data at time $t,\ t+k\in T$ is included in Table 1 :

Table 1. The output value $y_1 \sim y_7$ of system w at $t,\ t+k\in T$

y	y_1	y_2	y_3	y_4	y_5	y_6	y_7
y_t	1.63	1.78	1.06	1.43	1.81	1.94	1.09
$y_{t+k}^{\bar{F}}$	1.63	1.78	-	1.43	-	1.94	-

where $y_t=\{y_1,y_2,y_3,y_4,y_5,y_6,y_7\}$ in table 1 is characteristic value set which is composed of output value of data $m=\{x_1,x_2,x_3,x_4,x_5,x_6,x_7\}$ at time $t\in T$, $y_{t+k}^{\bar{F}}=\{y_1,y_2,y_4,y_6\}$ is characteristic value set which is composed of output value of data $m^{\bar{F}}=\{x_1,\ x_2,x_4,x_6\}$ at time $t+k\in T$, "-" denotes "air data".

Based on Table 1 and formula (17) in section 2 we can get $\rho^{\bar{F}}=\|y^{\bar{F}}\|/\|y\|=3.41/4.15=0.76<1$, or $\rho^{\bar{F}}<\rho=\|y\|/\|y\|=1$. By definitions 4, 5, $\bar{F}$-data circle $\mathrm{O}^{\bar{F}}$ generated by $m^{\bar{F}}$ and data circle O generated by m satisfy $\mathrm{O}^{\bar{F}}\subset\mathrm{O}$.Using theorem 4, we get that $\mathrm{O}^{\bar{F}}$ is internal-concentric circle of O.

On-line Query and Confirm of System w

The integral circuit (RC circuit) which is connected with x_3,x_5,x_7 in identification module network changes into "short circuit", so, it makes $y_3=y_5=y_7=0$. The results given in Table 1 are confirmed on-line. Repair network is startup in system w, the output of system w is normal, it meets $\bar{F}$-data restore guideline. In Fig.2 data circle O is an intermediate circle, $\bar{F}$-data circle $\mathrm{O}^{\bar{F}}$ is an internal concentric circle of O,$\mathrm{O}^{\bar{F}}$ is indicated by the thick solid line. F-data circle O^{F} is an outer concentric circle of O,O^{F} is indicated by the thick solid line. $\rho=1$ is the radius of O, $\rho^{\bar{F}}<1$ is the radius of $\mathrm{O}^{\bar{F}}$, $\rho^{F}>1$ is the radius of O^{F}.

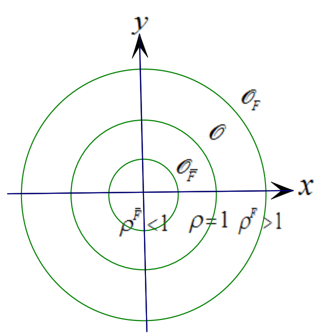


Fig. 2. Shows the intuitive graphical representation of data circle O, $\bar{F}$- data circle $\mathrm{O}^{\bar{F}}$, and F- data circle O^{F}

5 Discussion

The paper makes use of the structure and characteristics of P-sets to give the concepts of internal-data circle and outer-data circle. By using these concepts, from a theoretical point of view we recognize the nature of pathological output and use example to prove the usefulness of the results in this paper. P-sets are new theory and new method to study the dynamic information system.

P-sets can give research on the problems about application, the new results are:

- Characteristic of Information Hiding and its application
- Characteristic of Information Memory and its application
- Characteristic of Information Inheritance and its application
- Characteristic of Information Variation and its application
- Characteristic of Information Restore and its application
- Characteristic of Information Law Generation and its application
- Characteristic of Information Image Separation-hiding and its application
- Characteristic of Information Re-combination and its application
- Characteristic of Information Dependence Discovery and its application

Acknowledgements. This work was supported by the Foundation of Educational Department of Fujian Province, China (Grant No.JA09151, JB08177).

References

1. Shi, K.Q.: P-sets and its applications. An International Journal Advances in Systems Science and Application 9(2), 209–219 (2009)
2. Shi, K.Q.: P-sets. Journal of Shandong University (Natural Science) 43(11), 77–84 (2008) (in Chinese)
3. Shi, K.Q., Zhang, L.: Internal P-sets and data outer-recovery. Journal of Shandong University (Natural Science) 44(4), 8–14 (2009) (in Chinese)
4. Li, Y.Y.: Generation and Recovery of *F*-distortion Data. Journal of Jishou University (Natural Science) 31(3), 59-72 (2010) (in Chinese)
5. Yu, X.Q.: Identification and filtering of P-sets. Journal of Shandong University (Natural Science) 45(1), 94–98 (2010) (in Chinese)
6. Yu, X.Q.: Area characteristic of P-sets. Journal of Shandong University (Natural Science) 45(2), 99–102 (2010) (in Chinese)
7. Zhang, F., Chen, P., Zhang, L.: P-separation of P-sets and its applications. Journal of Shandong University (Natural Science) 45(3), 71–75 (2010) (in Chinese)
8. Shi K.Q., Li X.H.: Camouflaged information identification and its applications. An International Journal Advances in Systems Science and Applications, 10(2), 157-167(2010)

9. Lin H.K., Li Y.Y.: P-sets and its P-separation theorems. An International Journal Advances in Systems Science and Applications, 10(2), 209-215(2010)
10. Shi, K.Q.: Two direction S-rough sets. International Journal of Fuzzy Mathematics 2, 335–349 (2005)
11. Shi, K.Q., Yao, B.X.: Function S-rough sets and systems law identification. Sciences in China (F) 51(5), 499–510 (2008)

Convergent Properties of Arbitrary Order Neural Networks

Zhen-kun Huang and Chao Chen

School of Science, Jimei University, Xiamen 361021, Fujian, China
e-mail: hzk974226@jmu.edu.cn, zkhuangdoc@hotmail.com

Abstract. In this paper, we establish sufficient conditions for simplifying the checking that an isolated equilibrium point of autonomous neural network is exponentially stable. Meanwhile, these conditions are generalized to convergent properties of all equilibrium points of arbitrary order neural network in an open region of the state space. The explicit lower bound on the exponential convergence rate to an equilibrium point is also estimated.

Keywords: Neural networks, arbitrary order, equilibrium, exponential convergence.

1 Introduction

It is well known that traditional neural networks with first-order synaptic connections may encounter inevitable difficulties, i.e., these models are shown to have limitations such as limited capacity when used in pattern recognition problems and the case of optimization problems that can be solved using such models are also limited [1-2]. By incorporating arbitrary order synaptic connectivity into neural networks, authors in [1] proposed the following neural system:

$$\dot{x}_i(t) = -a_i(x_i)\left[b_i(x_i) + \sum_{k=1, i\in I_k}^{L} c_k \frac{m_i(k)}{d_i(x_i)} \prod_{j\in I_k} d_j^{m_j(k)}(x_j)\right], \quad i = 1, \cdots, N(1)$$

It is shown that high-order networks may improve dramatically their storage capacity in some degree and would increase the class of optimization problems that can be solved using neural networks [1-3]. Due to the advantage of high-order synaptic connectivity, there have been considerable works about stability analysis of high-order neural networks in the literature (e.g., see [4-8]). However, from existing reports, we can notice that most of the existing works of high-order neural networks [1-2], [4-8] have only fucus on the unique

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 593–601.
springerlink.com

equilibrium, periodic solution or almost periodic solution and their global attractivity; Seldom have been done for local stability and convergent estimate of arbitrary order neural networks. It is worth for us to further investigate convergent properties of of arbitrary order neural networks in local regions of state space.

2 Preliminaries

Consider the following autonomous differential equation

$$\dot{x} = F(x), \tag{2}$$

where $x \in R^N$. A constant vector $x^* \in R^N$ is said to be an equilibrium point of (2) if and only if $F(x^*) = 0$. Denote $|x| = \max_{i\in\mathcal{N}} |x_i|$ by the Euclidean norm of a vector $x \in R^N$.

Definition 1. *The equilibrium point x^* of (2) is said to be stable if there exists a constant $\delta(\varepsilon) > 0$ such that for every $t > 0$ and $|x(0) - x^*| < \delta(\varepsilon)$, we have $|x(t) - x^*| < \varepsilon$. The equilibrium point x^* is unstable if it's not stable.*

Definition 2. *The equilibrium point x^* of (2) is said to be locally exponentially stable if it's stable and there exist $M > 0$, $\lambda > 0$ and a neighborhood $O(x^*)$ such that $|x(t) - x^*| \le M|x(0) - x^*|e^{-\lambda t}$ $\forall t > 0$, where $x(0) \in O(x^*)$.*

Theorem 1. *([9]) Let x^* be an isolated equilibrium point of (2). If all the eigenvalues of Jacobian matrix $DF(x^*)$ have negative real parts, then x^* is asymptotically stable. If at least one eigenvalue has a positive real part, x^* is unstable.*

Theorem 2. *([10]) All the eigenvalue of a matrix $C = (c_{ij})_{N\times N} \in R^{N\times N}$ lie within the union of disks, in the complex plane, with centres $c_{ii}, i = 1, 2, \cdots N$ and radii*

$$r_i = \sum_{j=1 j\neq i}^{N} |c_{ij}|, \quad i = 1, 2, \cdots N.$$

Now consider the following general neural network

$$x_i^{'}(t) = -F_i(x_i(t)) + G_i(x_1(t), x_2(t), \cdots, x_N(t)) + I_i, \tag{3}$$

where $i \in \mathcal{N} := \{1, 2, \cdots, N\}$. Obviously, (3) includes many Hopfield-type neural networks [6,13] as its special cases.

Lemma 1. *([14]) Let $x^* = (x_1^*, x_2^*, \cdots, x_N^*)^T \in R^N$ be an isolated equilibrium point of (3). If*

$$F_i'(x_i^*) > \sum_{j=1}^{N} \left| \frac{\partial G_i(x^*)}{\partial x_j} \right|, \quad i = 1, 2, \cdots, N, \tag{4}$$

then x^* is locally exponentially stable and there exist $B > 0$, $\delta > 0$ and a neighborhood $O(x^*)$ such that

$$|x(t) - x^*| \leq B|x(0) - x^*|e^{-\delta t}, \quad \forall t > 0$$

where $x(0) \in O(x^*)$ and the convergent rate

$$\delta \geq \min_{i \in \mathcal{N}} \left\{ F_i'(x_i^*) - \sum_{j=1}^{N} \left| \frac{\partial G_i(x^*)}{\partial x_j} \right| \right\}.$$

Example 1. Consider Hopfield neural network with two neurons:

$$\begin{cases} \dot{x}_1(t) = -a_1 x_1 + t_{11} g_1(x_1) + t_{12} g_2(x_2) + I_1, \\ \dot{x}_2(t) = -a_2 x_2 + t_{21} g_1(x_1) + t_{22} g_2(x_2) + I_2. \end{cases} \tag{5}$$

Take $g_i(x) \equiv \tanh(x), i = 1, 2$. Design $a_1 = \frac{1}{\rho} + |t_{11}|g_1'(x_1^*) + |t_{12}|g_2'(x_2^*)$ and $a_2 = \frac{1}{\rho} + |t_{21}|g_1'(x_1^*) + |t_{22}|g_2'(x_2^*)$, where $\rho > 0$ and $x^* = (x_1^*, x_2^*)^T \in R^2$ is an isolated equilibrium point of (5). Then from Lemma 1, we can easily check that x^* is locally exponentially stable and the lower bound of exponential convergent rate is $\frac{1}{\rho}$.

3 Main Results

Consider the following arbitrary order networks [1]:

$$\dot{x}_i(t) = -F_i(x_i(t)) + \sum_{j=1}^{L} \prod_{k \in I_j} w_{ij}^k \Big[g_k(x_k(t)) \Big]^{d_k(j)} + I_i, \quad i = 1, 2, \cdots, N, \tag{6}$$

where $\{I_1, I_2, \cdots, I_L\}$ is a collection of L not-ordered subsets of $\mathcal{N}$, $d_k(j)$ is a nonnegative integer, $g_k(\cdot)$ is an activation function. Obviously, system (6) is a general high-order neural networks including [6] as its special cases. We make the following basic assumptions, $k = 1, 2, \cdots, N$:

$$\begin{cases} -g_k(x) = g_k(-x), g_k(0) = 0, \;\; 0 < F_k'(x) < \infty, \\ \\ 0 < g_k'(x) < g_k'(0), \;\; x g_k''(x) < 0 \;\text{ for }\; x \in R/\{0\}, \\ \\ \lim\limits_{x \to -\infty} g_k(x) = \check{g}_k, \;\; \lim\limits_{x \to +\infty} g_k(x) = \hat{g}_k. \end{cases}$$

Let x^* be an equilibrium point of (6), i.e.,

$$F_i(x_i^*) = \sum_{j=1}^{L} \prod_{k\in I_j} w_{ij}^k \Big[g_k(x_k^*)\Big]^{d_k(j)} + I_i, \quad i = 1, 2, \cdots, N.$$

Due to the strict monotonicity of $F_i(\cdot)$, we get

$$x_i^* = F_i^{-1}\left(\sum_{j=1}^{L} \prod_{k\in I_j} w_{ij}^k \Big[g_k(x_k^*)\Big]^{d_k(j)} + I_i\right), \quad i = 1, 2, \cdots, N$$

which leads to $|x_i^*| \le F_i^{-1}(\sum_{j=1}^{L} |w_{ij}^k| \prod_{k\in I_j} (g_k^\natural)^{d_k(j)} + |I_i|)$, where $g_k^\natural = \max\{|\check{g}_k|, \hat{g}_k\}$, $i, k = 1, 2, \cdots, N$. Hence, all equilibrium points of (6) lie in $V_\omega := V_\omega^1 \times V_\omega^2 \times \cdots \times V_\omega^N$, where

$$V_\omega^i := \left\{ x \in R \,\middle|\, |x| \le F_i^{-1}\left(\sum_{j=1}^{L} \prod_{k\in I_j} |w_{ij}^k| (g_k^\natural)^{d_k(j)} + |I_i|\right)\right\}, \; i = 1, 2, \cdots, N.$$

For convenience of discussing stability of all equilibrium points in an open region of the state space, we define

$$\mathcal{H}(\upsilon) = \Big\{ x \in R^N \Big| |g_k(x_k)| \ge \upsilon_k, \quad k = 1, 2, \cdots, N\Big\},$$

where the constant vector υ satisfies with $\upsilon = (\upsilon_1, \upsilon_2, \cdots, \upsilon_N)^T > 0$ and $\mathcal{H}(\upsilon) \cap V_\omega \neq \emptyset$. Due to the monotonicity of $g_k(\cdot)$, there exists a unique $\alpha_k > 0$ such that $g_k(\alpha_k) = \upsilon_k$ and $g_k(-\alpha_k) = -\upsilon_k$, i.e., $\alpha_k = g_k^{-1}(\upsilon_k), k = 1, 2, \cdots, N$. Moreover, if $|x_k| \ge \alpha_k$ for $k = 1, 2, \cdots, N$, then $|g_k(x_k)| \ge \upsilon_k$. Hence, $\mathcal{H}(\upsilon)$ is the union of 2^N disjoint regions:

$$\mathcal{H}(\upsilon) = \bigcup_{k=1}^{2^N} \Omega(\xi^{(k)}),$$

where

$$\Omega(\xi^{(k)}) = \left\{ x \in R^N \middle| |x_i| \ge \alpha_i, \; x_i\xi_i^{(k)} > 0, \; i = 1, 2, \cdots, N\right\}, \;\; \xi^{(k)} \in \Big\{-1, +1\Big\}^N.$$

Theorem 3. *Let x^* is an equilibrium point of (6) located in $\mathcal{H}(\upsilon)$. If*

$$\begin{aligned} F_i'(x_i^*) > \sum_{j=1}^{N} \Bigg[& \sum_{\ell=1, I_\ell\cap\{j\}\neq\emptyset}^{L} |w_{i\ell}^j| |g_j(x_j^*)|^{d_j(\ell)-1} d_j(\ell) \\ & \times \prod_{k\in I_\ell/\{j\}} |w_{i\ell}^k| |g_k(x_k^*)|^{d_k(\ell)}\Bigg] g_j'\Big(g_j^{-1}(\upsilon_j)\Big), \; i = 1, 2, \cdots, N \end{aligned} \quad (7)$$

then x^* is locally exponentially stable. Moreover, a lower bound of exponential convergence rate associated with x^* is given as

$$\delta \geq \min_i \left\{ F_i^{'}(x_i^*) - \sum_{j=1}^{N} \left[\sum_{\ell=1, I_\ell \cap \{j\} \neq \emptyset}^{L} |w_{i\ell}^j||g_j(x_j^*)|^{d_j(\ell)-1} d_j(\ell) \right.\right.$$
$$\left.\left. \times \prod_{k \in I_\ell / \{j\}} |w_{i\ell}^k||g_k(x_k^*)|^{d_k(\ell)} \right] g_j^{'}\Big(g_j^{-1}(\upsilon_j)\Big) \right\}.$$

Proof. From the definition of $g_k(\cdot)$, it follows that

$$g_i^{'}(u) < g_i^{'}(v) \text{ if } |u| > |v|$$

and

$$g_i^{'}(x_i) \leq g_i^{'}(\alpha_i) = g_i^{'}(g_i^{-1}(\upsilon_i)) \text{ if } x \in \Omega(\xi^{(k)}),$$

where $i = 1, 2, \cdots, N$, $k = 1, 2, \cdots, 2^N$. Hence, if x^* is an isolated equilibrium point located in $\Omega(\xi^{(k)})$, then $g_i^{'}(x_i^*) \leq g_i^{'}(g_i^{-1}(\upsilon_i))$, $i = 1, 2, \cdots, N$. From Lemma 1, if

$$F_i^{'}(x_i^*) > \sum_{j=1}^{N} \left\{ \sum_{\ell=1, I_\ell \cap \{j\} \neq \emptyset}^{L} |w_{i\ell}^j| d_j(\ell) g_j^{'}(x_j^*)|g_j(x_j^*)|^{d_j(\ell)-1} \right.$$
$$\left. \times \prod_{k \in I_\ell / \{j\}} |w_{i\ell}^k||g_k(x_k^*)|^{d_k(\ell)} \right\} \tag{8}$$

holds for each $i \in \mathcal{N}$, then x^* is locally exponentially stable. Obviously, (7) always implies (8) holds. Moreover, the lower bound of exponential convergence δ can be estimated as

$$\min_i \left\{ F_i^{'}(x_i^*) - \sum_{j=1}^{N} \left[\sum_{\ell=1, I_\ell \cap \{j\} \neq \emptyset}^{L} |w_{i\ell}^j||g_j(x_j^*)|^{d_j(\ell)-1} d_j(\ell) \right.\right.$$
$$\left.\left. \times \prod_{k \in I_\ell / \{j\}} |w_{i\ell}^k||g_k(x_k^*)|^{d_k(\ell)} \right] g_j^{'}\Big(g_j^{-1}(\upsilon_j)\Big) \right\}.$$

Since all the equilibrium points lie in V_ω, it's easy for us to have the following corollary:

Corollary 1. *Let x^* is an equilibrium point of (6) located in $\mathcal{H}(\upsilon)$. If*

$$F_i^{'}(x_i^*) > \sum_{j=1}^{N} \left[\sum_{\ell=1, I_\ell \cap \{j\} \neq \emptyset}^{L} |w_{i\ell}^j| \Big(\sup_{v \in V_\omega^j} |g_j(v)| \Big)^{d_j(\ell)-1} d_j(\ell) \right.$$
$$\left. \times \prod_{k \in I_\ell / \{j\}} |w_{i\ell}^k| \Big(\sup_{v \in V_\omega^k} |g_k(v)| \Big)^{d_k(\ell)} \right] g_j^{'}\Big(g_j^{-1}(\upsilon_j)\Big),$$

where $i = 1, 2, \cdots, N$, then x^* is locally exponentially stable. Moreover, a lower bound of exponential convergence rate δ is given as

$$\delta \geq \min_i \left\{ F_i'(x_i^*) - \sum_{j=1}^{N} \left[\sum_{\ell=1, I_\ell \cap \{j\} \neq \emptyset}^{L} |w_{i\ell}^j| \Big(\sup_{v \in V_\omega^j} |g_j(v)| \Big)^{d_j(\ell)-1} \right.\right.$$
$$\left.\left. \times\ d_j(\ell) \prod_{k \in I_\ell / \{j\}} |w_{i\ell}^k| \Big(\sup_{v \in V_\omega^k} |g_k(v)| \Big)^{d_k(\ell)} \right] g_j'\Big(g_j^{-1}(v_j)\Big) \right\}.$$

Corollary 2. *Let x^* is an equilibrium point of (6). If*

$$F_i'(x_i^*) = \sup_{v \in R} F_i'(v) > \sum_{j=1}^{N} \left[\sum_{\ell=1, I_\ell \cap \{j\} \neq \emptyset}^{L} |w_{i\ell}^j| (g_j^\natural)^{d_j(\ell)-1} d_j(\ell) \right.$$
$$\left. \times \prod_{k \in I_\ell / \{j\}} |w_{i\ell}^k| (g_k^\natural)^{d_k(\ell)} \right] g_j'(0), \tag{9}$$

then x^* is a unique equilibrium of (6) which is globally exponentially stable. δ can be estimated by

$$\delta \geq \min_i \left\{ \sup_{v \in R} F_i'(v) - \sum_{j=1}^{N} \left[\sum_{\ell=1, I_\ell \cap \{j\} \neq \emptyset}^{L} |w_{i\ell}^j| (g_j^\natural)^{d_j(\ell)-1} d_j(\ell) \right.\right.$$
$$\left.\left. \times \prod_{k \in I_\ell / \{j\}} |w_{i\ell}^k| (g_k^\natural)^{d_k(\ell)} \right] g_j'(0) \right\}.$$

Proof. From Corollary 1, we know that it's sufficient for us to verify the uniqueness of equilibrium point of (6). Suppose that there exist two equilibrium points x^*, x^{**} of (6). Then from (6), we have

$$F_i(x_i^*) - F_i(x_i^{**}) = \sum_{\ell=1}^{L} \prod_{k \in I_\ell} w_{i\ell}^k (g_k(x_k^*))^{d_k(\ell)} - \sum_{\ell=1}^{L} \prod_{k \in I_\ell} w_{i\ell}^k (g_k(x_k^{**}))^{d_k(\ell)}.$$

where $i = 1, 2, \cdots, N$. From the Lagrange mean value theorem, it follows that

$$F_i'(\xi_i)|x_i^* - x_i^{**}| \leq \sum_{j=1}^{N} \left[\sum_{\ell=1, I_\ell \cap \{j\} \neq \emptyset}^{L} d_j(\ell) (g_j^\natural)^{d_j(\ell)-1} |w_{i\ell}^j| \right.$$
$$\left. \times \prod_{k \in I_\ell / \{j\}} |w_{i\ell}^k| (g_k^\natural)^{d_k(\ell)} \right] g_j'(\xi_j^\ell) |x_j^* - x_j^{**}|$$

which lead to

$$|x^* - x^{**}| \leq \max_{i \in \mathcal{N}} \left\{ \frac{\sum_{j=1}^{N} \left[\sum_{\ell=1, I_\ell \cap \{j\} \neq \emptyset}^{L} d_j(\ell) (g_j^\natural)^{d_j(\ell)-1} \right.}{\sup_{v \in R} F_i'(v)} \right.$$
$$\left. \left. \times\ |w_{i\ell}^j| \prod_{k \in I_\ell / \{j\}} |w_{i\ell}^k| (g_k^\natural)^{d_k(\ell)} \right] g_j'(0) \right\} |x^* - x^{**}|,$$

where ξ_i lies between x_i^* and x_i^{**}, ξ_j^ℓ lies between x_j^* and x_j^{**}, $i, j \in \mathcal{N}$, $\ell = 1, 2, \cdots, L$. It follows from (9) that $x^* = x^{**}$.

Remark 1. Few works have been done for the local stability for arbitrary order neural networks. Our results not only provide simple criteria to check convergence of each equilibrium point of (6) but also provide an effective approach to estimate the location and convergence rate of equilibrium points in state space.

Example 2. Consider Hopfield-type neural network with second-order connections:

$$x_i^{'}(t) = -a_i x_i + \sum_{j=1}^{3} \prod_{k \in I_j} w_{ij}^k \Big(g_k(x_k)\Big)^{d_k(j)} + J_i, \;\; i = 1, 2 \tag{10}$$

where

$$\begin{aligned}
&a_1 = a_2 = 1, d_1(1) = d_2(2) = 1, d_1(3) = d_2(3) = 2,\\
&w_{11}^1 = 4, w_{12}^2 = 0, w_{13}^1 = 1, w_{13}^2 = 0.2, J_1 = 0.5,\\
&w_{21}^1 = 0, w_{22}^2 = 4, w_{23}^1 = 0.2, w_{23}^2 = 1, J_2 = 0.4,\\
&I_1 = \{1\}, I_2 = \{2\}, I_3 = \{1, 2\}, g_1(x) = g_2(x) = g(x) = \tanh(x).
\end{aligned}$$

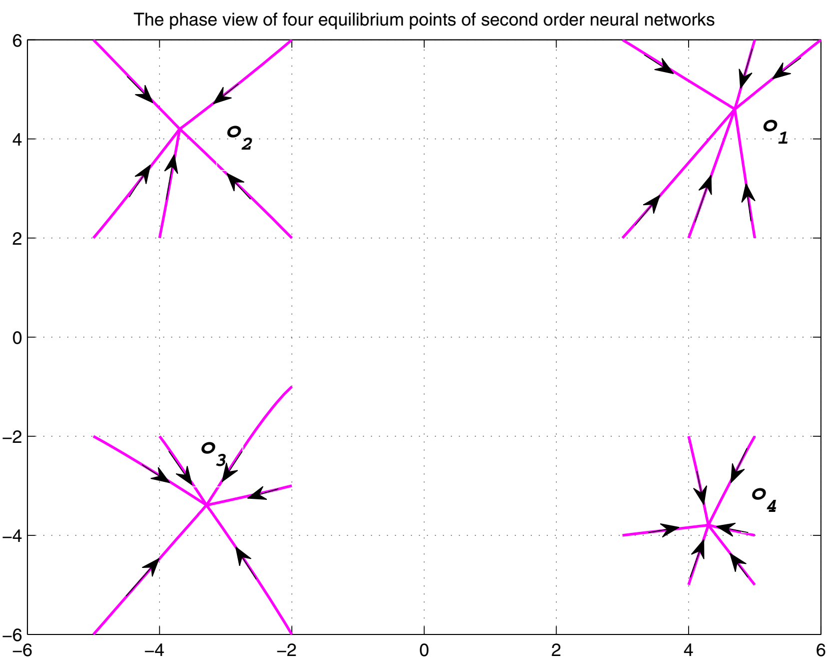


Fig. 1. The convergence of four equilibria of second order neural networks (10)

It is easy for us to get $V_\omega = [-4.7, 4.7] \times [-4.6, 4.6]$, $\alpha = (\alpha_1, \alpha_2)^T = (1.32, 1.32)^T$, $\upsilon = (\upsilon_1, \upsilon_2)^T = (0.8668, 0.8668)^T$, $H(\upsilon) = \bigcup_{k=1}^{4} \Omega(\xi^{(k)})$ and

$$\Omega(\xi^{(k)}) = \left\{ x \in R^2 \middle| |x_i| \geq 1.32,\ \ x_i \xi_i^{(k)} > 0,\ \ i = 1, 2 \right\},\ \ \xi^{(k)} \in \Big\{ -1, +1 \Big\}^2.$$

From computer numerical simulations, we can check that there exist four equilibrium points

$$o_1 = (4.699, 4.599)^T, o_2 = (-3.695, 4.199)^T,$$
$$o_3 = (-3.29, -3.392)^T, o_4 = (4.299, -3.796)^T.$$

By simple checking assumptions in Theorem 3, the four equilibrium points of (10) are local exponentially stable. The exponential convergent rate δ can also be estimated. For their phase view and the convergent dynamics of these equilibrium points, we can refer to Fig. 1.

Acknowledgements. The project is supported by the Foundation for Young Professors of Jimei University, the Scientific Research Foundation of Jimei University and the Foundation for Talented Youth with Innovation in Science and Technology of Fujian Province (2009J05009).

References

1. Dembo, A., Farotimi, O., Kailath, T.: High-order absolutely stable neural networks. IEEE Trans. Circuits Syst. 1(38), 57–65 (1991)
2. Kosmatopoulos, E.B., Christodoulou, M.A.: Structural properties of gradient recurrent high-order neural networks. IEEE Trans. Circuits Syst. II: analog and digital signal processing 42(9), 592–603 (1995)
3. Paretto, P., Niez, J.J.: Long term memory storage capacity of multiconnected neural networks. Biol. Cybern. 54, 53–63 (1986)
4. Lou, X., Cui, B.: Novel global stability criteria for high-order Hopfield-type neural networks with time-varying delays. J. Math. Anal. Appl. 330, 144–158 (2007)
5. Xu, B., Liu, X., Liao, X.: Global asymptotic stability of high-order Hopfield type neural networks with time delays. Computer and Mathematics with Applications 45, 1729–1737 (2003)
6. Xu, B., Liu, X., Liao, X.: Global exponential stability of high-order Hopfield type neural networks. Applied Mathematics and Computation 174, 98–116 (2006)
7. Yu, Y., Cai, M.: Existence and exponential stability of almost-periodic solutions for high-order Hopfield neural networks. Mathematical and Computer Modelling 47(9), 943–951 (2008)
8. Huang, Z., Xia, Y.H.: Exponential p-stability of second order Cohen-Grossberg neural networks with transmission delays and learning behavior. Simulation Modelling Practice and Theory 15, 622–634 (2007)

9. Guckenheimer, J., Holmes, P.: Nonlinear oscillations, dynamical systems, and bifurcations of vector fields. Springer, Berlin (1983)
10. Ralston, A.: A first course in numerical analysis. McGraw-Hill, New York (1965)
11. Hirsch, M.W., Smale, S.: Differential equations, dynamical systems, and linear algebra. Academic Press, London (1974)
12. Ren, F., Cao, J.: LMI-based criteria for stability of high-order neural networks with time-varying delay. Nonlinear Analysis: Real World Applications 7, 967–979 (2006)
13. Hopfield, J.J.: Neurons with graded response have collective computational properties like those of two-state neurons. Proc. Nat. Acad. Sci. 81, 3088–3092 (1984)

Analyzing and Predicting Interannual Variability in Nearshore Topographic Field on Large Scales by EOF and ANN

Jin-dong Wang[1] and Xiao-gong Sun[2]

[1] School of Marine and Atmospheric Sciences,
Stony Brook University, Stony Brook, New York, USA
e-mail: jindwang@ic.sunysb.edu
[2] Chinese Academy of Meteorological Sciences, Beijing, China
e-mail: sunxg@cams.cma.gov.cn

Abstract. Empirical Orthogonal Functions (EOF) and Artificial Neural Network (ANN) are performed for investigating and predicting interannual variability in nearshore topographic field around the Yellow River Delta. EOF and ANN are particularly effective at reproducing the observed topographic features around either modern or historical river mouths where the nearshore topography has experienced significantly intense interannual changes during the last thirty decades. In general, the coastal land around the modern river mouths has extended toward the sea and the seafloor has been elevated due to accumulation of a large amount of sediment transported by the modern Yellow River. On the other hand, the coast and the seafloor around the historical river mouths have been eroded quickly by tidal current, waves and storms due to lack of sediment supply. The observed spatial patterns of nearshore topography are well captured qualitatively and quantitatively by dominant eigenvectors of EOF. The EOF principal components indicating temporal variation in dominant eigenvectors are effectively fitted and predicted by using the observed river-related data as the input to ANN. The topographic fields either for the "missing" years or for the "future" years can then be estimated by linear combination of fixed dominant eigenvectors and fitted principal components. As a result, the fitting and predicting errors reach as low as 2.9% and 5.6%, respectively.

Keywords: EOF, ANN, interannual variability, nearshore topography, the yellow river delta.

1 Introduction

Nearshore topography, especially around the large rivers, usually experiences significant changes over temporal scales of days to decades and over spatial scales of meters to hundreds of kilometers due to sediment transport controlled mainly by river flow, tidal current, waves, storms, and precipitation. The understanding and prediction of nearshore topographic development are important for coastal

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 603–612.
springerlink.com

management and hydrodynamic simulation and prediction at the nearshore sea. The Yellow River, famous for a large amount of sediment discharge and frequent changes in the position of river channel and river mouth, leads to quick and intense variations in nearshore topography around the historical and modern river mouths. The relationship between the Yellow River input and variations in the coastal line [1-3], delta area [4], small-scale topography [5-7] has been studied a lot. However, little attention has been focused on quantitatively exploring interannual variability in nearshore topography around the whole Yellow River Delta and over decades. In order to acquire effective description and efficient prediction of topographic variation on large scales, a large number of multiyear bathymetry data at thousands of geographic points need to be simplified and the nonlinear relationship between the bathymetry data and the influencing factors need to be specified. Thus we come up with combination and application of EOF and ANN as an appropriate approach to solve this problem.

The central idea of EOF is to reduce dimensionality of the large data set interrelated in space and time but still to retain dominant variations of the data set by a few orthogonal functions. The structure of orthogonal functions is determined by the original data set and is not assumed a priori function form. Each of the orthogonal functions takes a certain proportion of total variability of the data set. The orthogonal functions with larger proportions can be seen as dominant functions containing the most information of the data set. In fact, each of the orthogonal functions can be represented by linear combination of an eigenvector and a principal component. The eigenvector indicates spatial pattern of the data set and can be used to identify physical meaning and relative influencing factors. The principal component suggests temporal variation in the corresponding spatial pattern. Thus the spatial and temporal information of a data set with a large number of variables can be efficiently represented by a few eigenvectors and principal components. EOF analysis was originally developed for meteorological application to data fields such as atmospheric pressure [8], precipitation [9], and sea surface temperature [10]. EOF was then introduced by coastal geologists and engineers into geomorphologic studies as an effective analytical tool to identify spatial and temporal patterns of beach topographic variability based on large data sets of beach profile surveys [11-14].

However, EOF itself can not be used to fit and predict temporal variability related with other influencing factors. ANN, known as a nonlinear analytical tool [15-16], has been applied to various areas [17-20] for processing and predicting nonlinear relationships. Thus ANN can be used in this study to identify and predict temporal variability of topographic field nonlinearly influenced by the factors like river parameters. Combination of advantages of EOF and ANN is capable of providing insights in characterization and prediction of spatial and temporal variability of large data sets.

The organization of this study is as follows. Section 2 describes the origin of the data sets and the methods of EOF and ANN. Section 3 describes the results of EOF analysis and the dominant patterns. Section 4 shows application of EOF and ANN in fitting and predicting topographic field. Section 5 is summary and conclusions.

2 Data and Methods

The bathymetry at the nearshore sea of the Yellow River Delta is measured in field surveys on an annual basis. The monitored area is about 6100 km^2, covering the sea located between N 37º22’24” and N 38º22’23”, E 118º28’34” and E 119º25’31”. Sixteen-year bathymetry data are used in this study, starting at 1976 and ending at 2000. The data at 1979, 1981-1984, 1995, 1997 are missed due to historical reasons. However, the results of the analysis have not shown obvious sensitivity to the missing data at in-between years. Annual river discharge and sediment discharge during 1976-2000 are involved in the analysis. The river-related data are obtained from observations at Lijin, a hydrologic station located at the lower reach of the Yellow River. In addition, interannual change in the direction of the river mouth is also considered in the analysis. The time series of river direction are obtained from interpretation of annual satellite images.

The bathymetry data are firstly interpolated into a uniform spatial grid through Kriging method. Each grid cell has the size of 1km×1km. The depth at each grid point represents one variable, which is assigned a sixteen-year time series. Each value in a time series can be seen as a sample of a certain variable. The mean value has been removed from the time series of each variable before further EOF analysis. The de-meaned data are then put into one matrix X with p rows and n columns. Each number in a row indicates a depth value at a grid point. Each number in a column represents a depth value at a certain year. Our goal is to decompose X (p×n) into the product of a matrix V (p×p) containing spatial vectors and a matrix Y (p×n) containing temporal vectors. The vectors in V and Y are orthogonal. In the following discussion, we will call spatial vectors as eigenvectors and temporal vectors as principal components.

$$X_{p\times n} = V_{p\times p} Y_{p\times n}. \tag{1}$$

Alternatively, an observation x_{ij} at a grid point i and a certain year j can be written as

$$x_{ij} = \sum_{k=1}^{p} v_{ik} y_{kj} = v_{i1} y_{1j} + v_{i2} y_{2j} + \ldots + v_{ip} y_{pj}. \tag{2}$$

The above goal can be achieved by eigen-decomposition of the covariance matrix XX’ (p×p).

$$XX' e_m = \lambda e_m, \tag{3}$$

where λ is an eigenvalue, e_m is an eigenvector. There should be p eigenvalues and p corresponding eigenvectors. Then the covariance matrix can be rewritten as:

$$XX' = V \Lambda V', \tag{4}$$

where Λ is the diagonal matrix containing all eigenvalues ordered by the magnitude of λ, V is the matrix containing all corresponding eigenvectors of XX’. Vectors of V are orthogonal:

$$VV' = V'V = I \,, \tag{5}$$

where I is a unit matrix.

If we assume that

$$Y = V'X \,. \tag{6}$$

Then we can testify that

$$YY' = V'XX'V = V'V\Lambda V'V = \Lambda \,. \tag{7}$$

So vectors of Y are orthogonal. Now both V and Y can be calculated from (4) and (6).

If the product of the first m (m<p) spatial vectors and the first m temporal vectors is seen as an estimate $\hat{x}_{ij}$ of the original data set, the total error can be estimated by Q:

$$Q = \sum_{i=1}^{p}\sum_{j=1}^{n}(x_{ij} - \hat{x}_{ij})^2 \quad , \tag{8}$$

where $\hat{x}_{ij} = \sum_{k=1}^{m} v_{ik} y_{kj}$

It can also be derived that Q is the sum of the last p-m eigenvalues of XX':

$$Q = \sum_{k=m+1}^{p} \lambda_k = \sum_{k=1}^{p} \lambda_k - \sum_{k=1}^{m} \lambda_k . \tag{9}$$

The sum of square values of the original data set S is:

$$S = \sum_{i=1}^{p}\sum_{j=1}^{n} x_{ij}^2 = \sum_{j=1}^{n} y_j V'Vy_j = \sum_{k=1}^{p} \lambda_k \ . \tag{10}$$

Thus the extent to which the first m eigenvectors and principal components can restore the original data set can be conveniently represented by the ratio of the sum of the first m eigenvalues to the sum of all eigenvalues:

$$G(m) = 1 - \frac{Q}{S} = \frac{\sum_{k=1}^{m} \lambda_k}{\sum_{k=1}^{p} \lambda_k} . \tag{11}$$

It should also be noted that total number of variables is usually in reality much larger than that of samples, i.e. p >> n. Eigen-decomposition of X'X (n x n) is much easier than that of XX' (p x p) due to much smaller dimension of X'X. It can be testified in linear algebra that the first m eigenvalues of X'X and XX' are the same and their eigenvectors have the relationship as follows:

$$e_m^{pp} = \lambda^{-1/2} \cdot X e_m^{nn} \,, \tag{12}$$

where e_m^{pp} and e_m^{nn} are the eigenvectors of XX' and X'X, respectively. Thus it is more efficient to calculate the eigenvalues and eigenvectors of X'X at first when p is larger than n. Then the eigenvectors of XX' can be easily obtained through (12).

A typical artificial neural network includes input layer, hidden layers, and output layer. Each layer is related with adjacent layers through weights and functions. The back-propagation algorithm, associated with supervised error-correction learning rule, is commonly used in training of ANN. The weights of the network are iteratively updated to approach the minimum of the error which is passed backward during training. The ANN in this study is built with one input layer with six input parameters, one hidden layer with three neurons, and one output layer with three output parameters. The six input parameters include: the annual river discharge, the annually-averaged accumulated river discharge, the annual sediment discharge, the annually-averaged accumulated sediment discharge, the averaged sediment concentration, and the direction of the river mouth. The three output parameters are the first three principal components from the EOF analysis. Log-Sigmoid activation functions are used for the hidden layer. Pureline functions are used for the output layer. The neural network with prescribed structure can be trained by using known input and output data. Appropriate weights are then determined through back-propagation error-correction algorithm. Once it has been trained completely, the neural network can be used to fit and predict output parameters if new input data are provided. Specific realization of this method is through MATLAB Neural Network Toolbox.

3 Spatial Pattern of Eigenvectors

From the EOF analysis, 92% of variation in topographic data can be restored by the first three eigenvectors and principal components. Study can now be focused on the first three eigenvectors rather than a large number of variables.

The first eigenvector contribute 79% to the total variance. Its spatial pattern is shown in Fig.1(a). Positive values of the first eigenvector have one local maximum center and are located around the modern Yellow River mouths. Negative values with three local minimum centers are related with the historical Yellow River mouths. The values close to zero distribute through outside sea, suggesting little interannual variation in submarine topography in these areas. In comparison, the difference in bathymetry data between 1976 and 1999 is shown in Fig.1(c). Positive values indicate deposition and elevation of sea floor. Negative values suggest erosion and sinking of sea floor. We can see clearly that deposition and erosion around the modern and historical Yellow River mouths matches very well with the spatial pattern of the first eigenvector.

Deposition and erosion around the Yellow River Delta is highly correlated with changes in sediment supply from the Yellow River [1, 4, 6]. From January 1964 to May 1976, the Yellow River went northward and trifurcated at the northern edge of the delta. A large amount of river sediment quickly accumulated near the river mouths and formed three centers of deposition with steep submarine slope. Since May 1976, the Yellow River have been running approximately eastward with

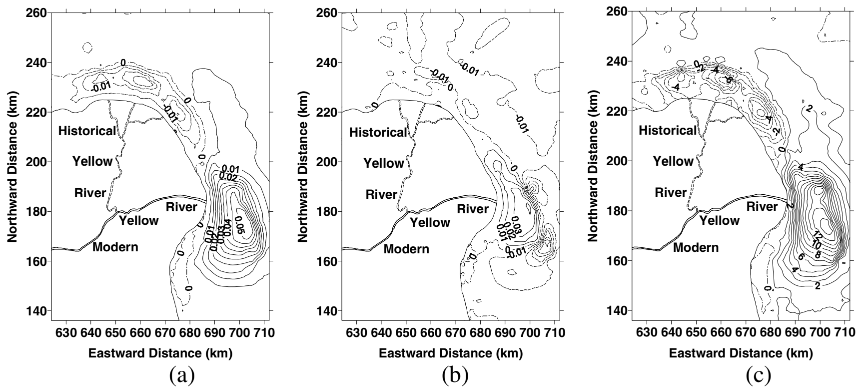


Fig. 1. Spatial pattern of: (a) the first eigenvector; (b) the second eigenvector; (c) deposition and erosion between 1976 and 1999

slight swing between the southeast and the northeast. The deposition center also moves to the modern river mouths. The old deposition centers are subject to quick erosion by ocean tides and waves due to lack of river sediment supply.

The second and third eigenvectors contribute to total variance with 10% and 3% respectively. All other eigenvectors contribute even less, representing small minor perturbations. Fig.1(b) shows spatial pattern of the second eigenvector. Topographic variation represented by the second eigenvector occurs mainly around the modern Yellow River mouths where submarine slope is further modified on the basis of first-eigenvector-represented topography in the direction of river-mouth extending. It should be noted that the lower ~20km reach of the modern Yellow River was artificially changed on a small scale from the direction toward the southeast to the direction toward the northeast in June, 1996. This change in river orientation is clearly shown by the second eigenvector which varies mainly in the directions of the southeast and northeast. The third eigenvector also modifies the submarine topography around the modern Yellow River mouths, which will not be discussed here for brief. But we can conclude that interannual variation in topography related to the modern river mouths is much more complicated and intense than that of the historical river mouths, and thus needs to be represented by more eigenvectors.

4 Fitting and Predicting Topographic Field

Principal components of EOF are actually temporal weights on spatial eigenvectors, representing different response of main features in different time. As discussed above, the first eigenvector shows sea floor erosion around the historical river mouths and deposition around the modern river mouths on a large spatial scale. The first principal component shows a decreasing trend of interannual variation (see Fig.2(a)), suggesting that the rate of either deposition or erosion also

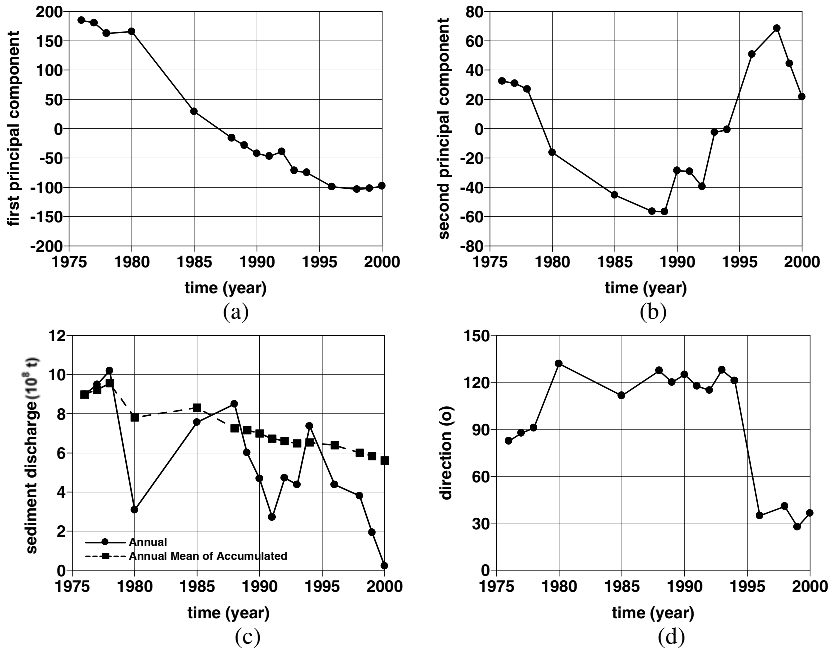


Fig. 2. Temporal variation of: (a) the first principal component; (b) the second principal component; (c) sediment discharge; (d) river mouth orientation

have decreased since 1976. Both erosion and deposition are related with the amount of sediment supply from the Yellow River. The time series of annual sediment discharge show in general a descending tendency (see Fig.2(c)). Since sediment amount in preceding years also have an effect on topographic variation in current year, it is summed up and annually averaged. The calculation result is defined as an annual mean of accumulated sediment discharge. The correlation coefficient between the first principal component and the annual mean of accumulated sediment discharge is as high as 0.92. From Table 1, we can see that the first principal component is also highly correlated with water discharge. Thus water and sediment supply from the Yellow River can be seen as dominant factors influencing topographic variation.

The second principal component decreases from 1976 to 1988 and then increases from 1988 to 2000. It may indicate that change in river mouth orientation (see Fig.2(d)) affects submarine topography in certain directions. The second principal component is negatively correlated with the river mouth orientation with correlation coefficient of -0.82. Thus river mouth orientation is also an important topographic influencing factor. The third principal component which is not shown here for brief has significant correlation with the annual sediment concentration which is an indicator of the ratio between water discharge and sediment discharge. This suggests that the water-sediment ratio also affects some details of topographic

Table 1. Correlation coefficients between principal components and river parameters

River Parameters	First PC	Second PC	Third PC
Annual water discharge	0.64	-0.22	0.29
Annual mean of accumulated water	0.79	-0.10	0.17
Annual sediment discharge	0.63	-0.11	0.23
Annual mean of accumulated sediment	0.93	-0.07	0.21
Annual sediment concentration	0.13	-0.06	0.50
River mouth orientation	0.29	-0.82	-0.17

variation. Different combination of river parameters has different relationship with principal components. Thus temporal variability in topography nonlinearly correlates with river input, which will be further specified by using ANN in the next discussion.

The artificial neural network discussed in section 2 is now utilized as a nonlinear modeling system to estimate the relationship between river parameters and principal components. In order to verify the EOF-ANN modeling system, we designed two experiments. One experiment is for fitting topographic field in the missing years. The other is for predicting topographic field in the future years. In the first experiment, we use the fifteen-year bathymetry data from 1976 to 2000 excluding 1992 for EOF analysis. We then train ANN by using river-related data as input and the first three principal components as output. Three values of principal component for 1992 can be estimated by trained ANN if the river data of 1992 are provided. The products of the principal component values and the corresponding eigenvectors can be summed up and added with the mean bathymetry to estimate the bathymetry in 1992. Comparison between observational values and model-estimated values indicates that they are consistent with each other very well (see Fig.3(a)). The mean error of all grid points is as low as 2.9%. Thus EOF combined with ANN is an effective tool to fit the bathymetry data in the missing years.

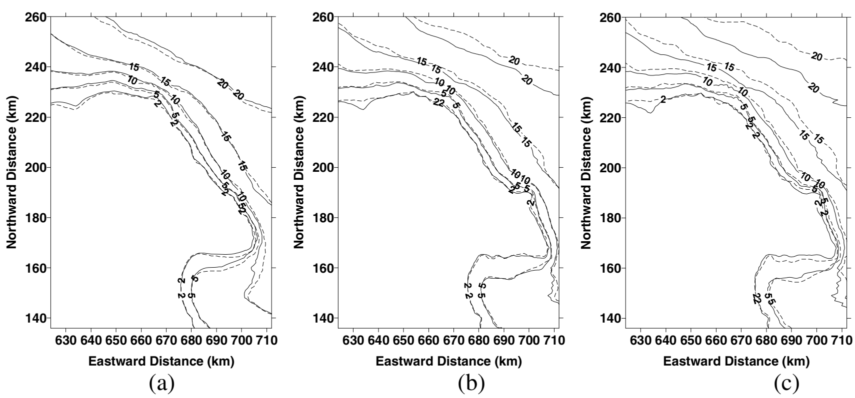


Fig. 3. Comparison between observational topographic contours and: (a) fitted topographic contours in 1992; (b) predicted topographic contours in 1999; (c) predicted topographic contours in 2000. Solid lines are observed; dashed lines are fitted or predicted.

In the second experiment, the fourteen-year bathymetry data from 1976 to 1998 are used for EOF and ANN analysis. The bathymetry data of 1999 and 2000 are then predicted by using the same procedures as the first experiment. The results are shown in Fig.3(b) and Fig.3(c). The mean error of prediction for 1999 and 2000 is 5.6%, and 11.9% respectively. Prediction of topographic field is generally accurate, but not as good as fitting experiment. Actually forecast of realistic variables is always harder than hindcast of realistic variables in natural sciences. In addition, nearshore prediction is better than off-shore. The errors are mainly related with the depth values located at the northeast corner of the study area, where interannual topographic variation may be rarely affected by water and sediment supply from the Yellow River. More precise prediction of off-shore topography needs specification of other influencing factors like currents and waves, which should be discussed in future study.

5 Summary and Conclusion

We have applied EOF to analyze the characteristics of spatial and temporal variation in submarine topography in the nearshore sea of the Yellow River Delta. EOF not only simplifies a large amount of variables into a few eigenvectors but also captures the dominant features of spatial and temporal variation in topographic data. The first eigenvector represents the processes of deposition and erosion in the nearshore sea around the modern and historical river mouths, which is mainly controlled by water and sediment supply from the Yellow River. The second eigenvector shows the effects of river mouth orientation on topographic variation around the modern river mouths. ANN is built as an effective nonlinear tool to evaluate the relationship between river input and principal components. EOF, in combination with ANN, is effectively used for fitting and predicting topographic field in the missing years and future years. Application of EOF and ANN model can also be extended to simplification, fitting and prediction of other large data sets featured by spatial and temporal variability.

Acknowledgements. This work was supported by the National Basic Research Program of China (973 Program) (No. 2002CB42404).

References

1. Pang, J.Z., Jiang, M.X., Li, F.L.: Changes and development trend of runoff sediment dis-charge and coastline of the Yellow River Estuary. Transactions of Oceanology and Limnology 4, 1–6 (2000)
2. Chang, J., Liu, G.H., Liu, Q.S.: Dynamic monitoring of coastline in the Yellow River Delta by remote sensing. Geo-Information Science 6(1), 94–98 (2004)
3. Huang, H.J., Fan, H.: Monitoring changes of neashore zones in the Yellow River Delta since 1976. Oceanologia et Limnologia Sinica 35(4), 306–314 (2004)
4. Sun, X.G., Yang, Z.S.: Prediction of modern Huanghe River Delta increment by using the sediment discharge. Oceanologia et Limnologia Sinica 26(1), 76–82 (1995)

5. Zhou, Y.Q.: Erosion process and main features of underwater coast slope of the northern coast in the Yellow River Delta. Marine Geology & Quaternary Geology 18(3), 79–85 (1998)
6. Liu, Y., Li, G.X., Deng, S.G., et al.: Evolution of erosion and accumulation in the aban-doned subaqueous delta lobe of the Yellow River. Marine Geology & Quaternary Geology 22(3), 27–34 (2002)
7. Sun, X.G., Yang, Z.S., Chen, Z.R.: Quantitative analysis of sediment erosion and accumu-lation in the nearshore sea of the modern Yellow River mouth. Acta Oceanologica Sinica 15(1), 129–136 (1993)
8. Craddock, J.M., Flood, C.R.: Eigenvectors for representing the 500mb geopotential surface over the Northern Hemisphere. Quarterly Journal of the Royal Meteorological Society 95, 576–593 (1969)
9. Harnack, R.P., Lanzante, J.R.: Specification of United States seasonal precipitation. Monthly Weather Review 113, 319–325 (1985)
10. Kutzbach, J.E.: Large-scale features of monthly mean Northern Hemisphere anomaly maps of sea level pressure. Monthly Weather Review 98, 708–716 (1970)
11. Winant, C.D., Inman, D.L., Nordstrom, C.E.: Description of seasonal beach changes using empirical eigenfunctions. Journal of Geophysical Research 80(15), 1979–1986 (1975)
12. Aubrey, D.G.: Seasonal patterns of onshore/offshore sediment movement. Journal of Geophysical Research 84(C10), 6347–6354 (1979)
13. Wijnberg, K.M., Terwindt, J.H.J.: Extracting decadal morphological behaviour from high-resolution, long-term bathymetric surveys along the Holland coast using eigenfunction analysis. Marine Geology 126, 301–330 (1995)
14. Larson, M., Hanson, H., Kraus, N.C., et al.: Short- and long-term responses of beach fills determined by EOF analysis. Journal of Waterways, Port, Coastal & Ocean Engineering 125(6), 285–293 (1999)
15. Hornik, K., Stinchcombe, M., White, H.: Multilayer feedforward networks are universal approximators. Neural Networks 2, 359–366 (1989)
16. Chen, S., Billings, S.A., Grant, P.M.: Non-linear system identification using neural net-works. International Journal of Control 51, 1191–1214 (1990)
17. Trippi, R.R., Turban, E.: Neural Networks in Finance and Investing, Probus, p. 513 (1993)
18. Lek, S., Delacoste, M., Baran, P., et al.: Application of neural networks to modelling nonlinear relationships in ecology. Ecological Modelling 90, 39–52 (1996)
19. Hsieh, W.W., Tang, B.Y.: Applying neural network models to prediction and data analy-sis in meteorology and oceanography. Bulletin of the American Meteorological Society 79(9), 1855–1870 (1998)
20. Maier, H.R., Dandy, G.C.: Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environmental Modelling & Software 15, 101–124 (2000)

f-Model Generated by P-Set

Yu-feng Qiu and Bao-hui Chen

School of mathematics and computer Sciences, Longyan University, Longyan, Fujian 364000, China
e-mail: qiuyffjly@163.com

Abstract. P-set is a set pair, which is composed of the internal and outer packet set together, and it has dynamic characteristics. Based on the theory of P-set, several concepts, such as f-model, the generation of f-t order unilateral dynamic model and so on are proposed; order relation theorem of f-unilateral dynamic model and dynamic separation theorem are put forward. The dynamic characteristic of f-model depends on the attribute supplement on α. Using those discussions, the generation principle of f-model and its application are given in the end.

Keywords: P-set, f-model, f-model Theorem, Unilateral Dynamic Characteristic.

1 Introduction

The prerequisite for all kinds of decision-making is that all information indexes are accurately predicted, but usually an information system is a dynamic and complex system in which there exist non-linear, time-variability and uncertainty, it is very difficult to predict those indexes accurately. In practice, people often use different mathematical models to predict kinds of information indexes. In 2008, Prof. Shi introduced dynamic characteristic into the general set (Cantor set) X, improved it and originated Packet sets $(X^{\bar{F}}, X^{F})$, written as P-set. P-set is a set pair, which is composed by internal and outer packet sets, and it has dynamic characteristics. Using the simple facts and the features of P-set, several concepts, such as f-model, the generation of f-t order unilateral dynamic model and so on, are proposed; the order relation theorem of f-unilateral dynamic model and dynamic separation theorem are put forward. The dynamic characteristic of f-model depends on the attribute supplement on α. Using those discussions, the generation principle of f-model and its application are given in the end. The discussion in paper shows that P-set theory [1,2] is an important theory in the research

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 613–620.
springerlink.com

of dynamic information system in analysis, modeling, forecasting, decision-making and control. Modeling to dynamic information system by P-set is a new research direction.

In order to facilitate the discussion, and to accept the following results easily and also to maintain the integrity of this article, the P-set and its structure is simply introduced into the next section as theoretical basis and preliminary.

2 *P*-Set and Its Structure

Assumption: *X is a finite general set on U, U is a finite element universe and V is a finite attribute universe.*

Definition 2.1. *Given a general set $X = \{x_1, x_2, \cdots, x_m\} \subset U$, and $\alpha = \{\alpha_1, \alpha_2, \cdots, \alpha_k\} \subset V$ is the attribute set of X, $X^{\bar{F}}$ is called internal packet sets of X, called internal P-set for short, moreover*

$$X^{\bar{F}} = X - X^{-}, \tag{1}$$

and X^{-} is called $\bar{F}$-element deleted set of X, moreover

$$X^{-} = \{x | x \in X, \bar{f}(x) = u \bar{\in} X, \bar{f} \in \bar{F}\}, \tag{2}$$

if the attribute set α^{F} of $X^{\bar{F}}$ satisfies

$$\alpha^{F} = \alpha \cup \{\alpha' | f(\beta) = \alpha' \in \alpha, f \in F\}, \tag{3}$$

where $\beta \in V$, $\beta \bar{\in} \alpha$, $f \in F$ turns β into $f(\beta) = \alpha' \in \alpha$.

Definition 2.2. *Given a general set $X = \{x_1, x_2, \cdots, x_m\} \subset U$, and $\alpha = \{\alpha_1, \alpha_2, \cdots, \alpha_k\} \subset V$ is the attribute set of X, X^{F} is called outer packet sets of X, called outer P-set for short, moreover*

$$X^{F} = X \cup X^{+}, \tag{4}$$

and X^{+} is called F-element supplemented set, moreover

$$X^{+} = \{u | u \in U, u \bar{\in} X, f(u) \in x' \in X, f \in F\}, \tag{5}$$

if the attribute set $\alpha^{\bar{F}}$ of X^{F} satisfies

$$\alpha^{\bar{F}} = \alpha - \{\beta_i | \bar{f}(\alpha_i) = \beta_i, \bar{f} \in \bar{F}\}, \tag{6}$$

where $\alpha_i \in \alpha$, $\bar{f} \in \bar{F}$ turns α_i into $\bar{f}(\alpha_i) = \beta_i \bar{\in} \alpha$.

Definition 2.3. *The set pair which is composed of internal P-set $X^{\bar{F}}$ and outer P-set X^{F} is called P-set (packet sets) generated by the general set X, called P-set for short, written as*

$$(X^{\bar{F}}, X^F), \tag{7}$$

where the general set X is called the ground set of $(X^{\bar{F}}, X^F)$.

Direct explanation of the name of P-set

Since the existence of element transfer $\bar{f} \in \bar{F}$, the element number of general set X in (6) decreases, X generates $X^{\bar{F}}$, and $X^{\bar{F}}$ is packed in X. Since the existence of element transfer $f \in F$, the element number of general set X in (9) increases, X generates X^F, and X^F is packed outside X. $X^{\bar{F}}$ and X^F are in the state of moving. $X^{\bar{F}}$ packed in X and X^F packed outside X together compose P-set $(X^{\bar{F}}, X^F)$.

The generation principle of P-sets

While some elements in general set X are transferred out of X, and some out of the set X are transferred into$X$, the general set X generates P-set $(X^{\bar{F}}, X^F)$, which has dynamic characteristic. The existence of $(X^{\bar{F}}$,$X^F)$depends on X, but it has no relation to how many elements are transferred out of X or into X.

3 f-Model and Generation of f-t Order Unilateral Dynamic Model

Definition 3.1. *Given a subset $X = \{x_1, x_2, \cdots, x_m\}$ of U, and every element x_i in X has characteristic data sequences $x_i^{(0)} = (x^{(0)}(1)_i, x^{(0)}(2)_i, \cdots, x^{(0)}(n)_i)$, $\forall x^{(0)}(k)_i \in R^+$, $k = 1, 2, \cdots, n$. $x^{(0)}$ is called the broken line model generated by X, if*

$$x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)), \tag{8}$$

and for any k, $x^{(0)}(k)$ satisfies

$$x^{(0)}(k) = \sum_{j=1}^{k} \sum_{i=1}^{m} x^{(0)}(j)_i, \tag{9}$$

here $x^{(0)}(k)$ is the feature value of X at the point $k = 1, 2, \cdots, n$, $x^{(0)}(k) \in R^+$. Obviously, $x^{(0)}$ has the increasing feature.

Definition 3.2. *$p(k)$ is called f-model generated by X, if*

$$p(k) = (1 - e^a)(x^{(0)}(1) - \frac{c}{a})e^{-ak}, \tag{10}$$

here a and c are parameters to be determined, seen Refs.[3,4].

Definition 3.3. *Given a subset $X = \{x_1, x_2, \cdots, x_m\}$ of U and the subset $X^{\bar{F}} = \{x_1, x_2, \cdots, x_t\}$ of X, $t \in N^+$, $t < m$, and every element x_i in*

$X^{\bar{F}}$ has characteristic data sequences $x_i^{(0)} = (x^{(0)}(1)_i, x^{(0)}(2)_i, \cdots, x^{(0)}(n)_i)$, $\forall x^{(0)}(k)_i \in R^+$, $k = 1, 2, \cdots, n$. $y^{(0)}$ is called the t order unilateral dynamic broken line model generated by X, if

$$y^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right), \tag{11}$$

and for any k, $x^{(0)}(k)$ satisfies

$$x^{(0)}(k) = \sum_{j=1}^{k}\sum_{i=1}^{t} x^{(0)}(j)_i, \tag{12}$$

here $y^{(0)}$ is the feature value of $X^{\bar{F}}$ at point $k = 1, 2, \cdots, n$. $x^{(0)}(k) \in R^+$. $y^{(0)}$ has the increasing feature.

Definition 3.4. *$p(k)^f$ is called f- t order unilateral dynamic model, if*

$$p(k)^f = \left(1 - e^b\right)\left(x^{(0)}(1) - \frac{d}{b}\right)e^{-bk}, \tag{13}$$

where $\alpha = \{\alpha_1, \alpha_2, \cdots, \alpha_\lambda\}$ is the attribute set of X and $\exists \beta \bar{\in} \alpha$, $f(\beta) = \alpha_{\lambda+1} \in \alpha$; $\alpha^f = \alpha \cup \{f(\beta)\} = \{\alpha_1, \alpha_2, \cdots, \alpha_\lambda, \alpha_{\lambda+1}\}$ is the attribute set of $X^{\bar{F}}$. $p(k)^f$ is the f-model generated by $X^{\bar{F}}$. b and d are the parameters to be determined, seen Refs.[3, 4].

Definition 3.5. *f-model $p(k)^f$ generated by $X^{\bar{F}}$ is the f-source model generated by X, written as $p(k)^*$, if*

$$\mathrm{card}(X^{\bar{F}}) = 1, \tag{14}$$

here card(X)=m, $m > 1$. Obviously, X has m f-source models.

By definitions 3.1-3.5, the following propositions can be obtained.

Proposition 3.1. *If α and α^f are attribute sets of X and $X^{\bar{F}}$ respectively, then*

$$\mathrm{card}(\alpha) \le \mathrm{card}(\alpha^f). \tag{15}$$

Proposition 3.2. *$p(k)^f$ and $p(k)$ have the same model feature, vice versa.*

Proposition 3.3. *f-source model $p(k)^*$ satisfies*

$$p(k)^* \le p(k)^f. \tag{16}$$

Proposition 3.4. *The f-t order unilateral dynamic model $p(k)^f$ satisfies*

$$p(k)^f \le p(k). \tag{17}$$

Theorem 3.1. *(The existence theorem of f-model) If* $X = \{x_1, x_2, \cdots, x_m\} \subset U$ *and each element* x_i *has characteristic data sequences* $x_i^{(0)} = (x^{(0)}(1)_i, x^{(0)}(2)_i, \cdots, x^{(0)}(n)_i)$, $\forall x^{(0)}(k)_i \in R^+$, $k = 1, 2, \cdots, n$, *and* $x^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n))$, *for any* k, $x^{(0)}(k) = \sum_{j=1}^{k} \sum_{i=1}^{t} x^{(0)}(j)_i$, *then there exists the f-model generated by X.*

It can be directly gotten from Refs [3,4], so its proof is omitted here.

From theorem 3.1, we can easily obtain the following two theorems.

Theorem 3.2. *(The existence theorem of f-t order unilateral dynamic model) if* $X^{\bar{F}} \subseteq X$ *, then the necessary and sufficient condition of the existence of* $p(k)^f$ *is that*

$$\mathrm{card}(X) \geq 1. \tag{18}$$

Theorem 3.3. *(The existence theorem of f-source model) if* $p(k)^*$ *is the f-source model of X, then the necessary and sufficient condition of the existence of* $p(k)^*$ *is that*

$$\mathrm{card}(X \cap X^{\bar{F}}) = 1. \tag{19}$$

The proofs can be directly obtained from above theorems and definition, so they are omitted.

Theorem 3.4. *(The order relation theorem of f-unilateral dynamic model) Let* $\pi = \left\{ P(k)_j^f | j = 1, \cdots, t \right\}$ *be the family composed of f-unilateral dynamic model of* $P(k)^f$*, and let* α_j^f *be the attribute set of* $P(k)_j^f$*. If*

$$\alpha \subseteq \alpha_1^f \subseteq \alpha_2^f \subseteq \cdots \subseteq \alpha_t^f, \tag{20}$$

then

$$P(k)_t^f \leq P(k)_{t-1}^f \leq \cdots \leq P(k)_1^f \leq P(k), \tag{21}$$

here α is the attribute set of X.

The proof can be obtained from properties of P-set, so omitted here.

Theorem 3.5. (The dynamic separation theorem of f-model) If $P(k)$ is the f-model generated by X, then $P(k)$ can be separated into a finite number of f-unilateral dynamic models $P(k)_j^f$, and their attribute sets satisfy

$$\alpha \subseteq \alpha_j^F. \tag{22}$$

Discovery principle of f-unilateral dynamic model

The attribute sets of f-unilateral dynamic model $P(k)_j^f$ satisfy $\alpha_j^f \subseteq \alpha_k^f$, then f-unilateral dynamic model $P(k)_k^f$ which has the attribute set α_k^f can be discovered in $P(k)_j^f$, and its existence has nothing to do with the size of $\mathrm{card}(\alpha_j^f)$, $\alpha_j^f \neq \emptyset$.

4 Application in Information System Prediction

Assumption: For simplification, but also without loss of generality, we suppose that the subset X is $\{x_1,\ x_2, x_3\ , x_4\}$. X is the subsystem of the economic information system $\mathbb{N}$, its attribute set is α, $\alpha = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$, the specific meaning of each α_i is omitted here. $x_j^{(0)}$ is the characteristic value sequence of x_j in some interval, $j = 1, 2, 3, 4$, it is shown in the following table.

Table 1. The state values of the subsystem $[u]$

	$x^{(0)}(1)_i$	$x^{(0)}(2)_i$	$x^{(0)}(3)_i$	$x^{(0)}(4)_i$	$x^{(0)}(5)_i$
$x_1^{(0)}$	2.15	2.06	2.02	2.45	2.90
$x_2^{(0)}$	2.78	3.03	3.08	2.92	2.31
$x_3^{(0)}$	3.01	2.87	2.59	2.35	3.08
$x_4^{(0)}$	2.08	1.98	3.01	3.23	3.07

Using $x^{(0)}(k) = \sum_{j=1}^{k}\sum_{i=1}^{4} x^{(0)}(j)_i$, $k = 1, 2, 3, 4$, we can get the broken line model:

$$\begin{aligned} x^{(0)} &= (x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), x^{(0)}(4), x^{(0)}(5)) \\ &= (10.02, 19.96, 30.66, 41.61, 52.97). \end{aligned} \tag{23}$$

Using data generation theory [3] and parametric equation, we obtain a generated model $P(k)$, moreover,

$$P(k) = (1 - e^a)(x^{(0)}(1) - \frac{c}{a})e^{-ak} = 9.6622e^{0.04171k}. \tag{24}$$

The above model (24) can be as a model of the subsystem of economic information system $\mathbb{N}$ in an interval at the time of the attribute set α unchanged, in other words, the motion state law [5-7] shown by the subsystem is the expression (24), we can use this expression predict the law of the subsystem at some interval.

In fact, the running status of economic system is changeable, some unknown attribute $\beta \bar{\in} \alpha$ attacks the attribute set α, in other words, some risk attributes break into α, i.e., $\beta_1, \beta_2 \bar{\in} \alpha$, $f(\beta_1) = \alpha_i$, $f(\beta_2) = \alpha_j$ and moreover $f(\beta_1)$, $f(\beta_2) \in \alpha$. α changes into $\alpha^f = \alpha \cup \{f(\beta_1), f(\beta_2)\} = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_i, \alpha_j\}$, the attribute invasion cause the changes of f-model. Actually, f-unilateral dynamic model appears at this moment, we have to use this f-unilateral dynamic model, when we forecast running law of the economy. If there exist the attacked risk attribute, then α changes into α^f , i.e., $\alpha^f = \alpha \cup \{f(\beta)\} = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_i, \alpha_j\}$. Under this condition, $X = X^{\bar{F}} = \{x_2, x_4\}$, the characteristic data of $X^{\bar{F}}$ is shown in Table 2, where, $i = 2, 4$.

Table 2. The state values of the subsystem $[u]^f$

	$x^{(0)}(1)_i$	$x^{(0)}(2)_i$	$x^{(0)}(3)_i$	$x^{(0)}(4)_i$	$x^{(0)}(5)_i$
$x_2^{(0)}$	2.78	3.03	3.08	2.92	2.31
$x_4^{(0)}$	2.08	1.98	3.01	3.23	3.07

Using the data in Table 2 and referring to the calculation process of the generated model $P(k)$, we obtain f- unilateral dynamic model of X, that is f- model $P(k)^f$ of $X^{\bar{F}}$,

$$p(k)^f = (1 - e^b)(x^{(0)}(1) - \frac{d}{b})e^{-bk} = 5.3886e^{0.01938k}. \tag{25}$$

The expression (25) tell us that we estimate running status of the system when the attribute invasion took place, we have to use f-unilateral dynamic model.

5 Conclusion

Refs [1, 2] originated Packet sets theory, which has the dynamic characteristic. P-set gives a theoretical support for searching the motion system model. The function on $[a, b]$ is just a model on $[a, b]$. P-set improved the general set and P-set is the generalization of the general set. Using P-set theory, this paper gives the research of changes of system state model in the case of attribute F- attack, and gives applications. The application presented in this paper just show how to predict and how to deal with some special cases in venture capital system and economic information system, the other problems can also be dealt with by P-set theory, we will give more discussions in other papers.

Acknowledgements. This work is supported by the Open Project of Brain-like Key Laboratory Research of Fujian Province (BLISSOS2010105), the Natural Science Foundation of Fujian Province of China (No. 2009J01293) and the A-kind Scientific Research Item of Education Department of Fujian Province of China (No. JA09151, JA09230).

References

1. Shi, K.: P-sets and its applications. An International Journal Advances in Systems Science and Applications 2, 168–178 (2009)
2. Shi, K.: P-set. Journal of Shandong University(Natural Science) 11, 77–84 (2008)
3. Zhang, L., Ren, X.: A F-modal generated by function one direction S-rough set. Journal of Shandong University (Natural Science) 8, 77–81, 89 (2008)

4. Deng, J.: Introduction to grey system theory. International Journal of Grey System 1, 1–24 (1989)
5. Zhang, L., Qiu, Y., Ren, X.: Rough law generation and its separation. Journal of Shandong University(Natural Sciences) 3, 58–63 (2008)
6. Zhang, L., Kaiquan, S.: Security transmission and recognition of F-Knowledge. Journal of Systems Engineering and Electronics 4, 877–882 (2009)
7. Shi, K.: S-rough sets and its applications in diagnosis-recognition for disease. In: IEEE Proceedings of the First International Conference on Machine Learning and Cybernetics, vol. 1, pp. 50–54 (2002)

Mining Fuzzy Association Rules by Using Nonlinear Particle Swarm Optimization

Guo-rong cai[1,2], Shao-zi Li[2], and Shui-li Chen[1,*]

[1] School of Science, Jimei University, Xiamen, 361021, China
e-mail: kiser@jmu.edu.cn, sgzx@jmu.edu.cn
[2] Deparment of Cognitive Science, Xiamen University, Xiamen, 360005, China
e-mail: southy@xmu.edu.cn

Abstract. This paper presents a fuzzy association rules mining algorithm by using nonlinear particle swarm optimization (NPSO) to determine appropriate fuzzy membership functions that cover the domains of quantitative attributes. Experiments conducted on the United States census demonstrated the feasibility and the efficiency of the proposed.

Keywords: Fuzzy association rule, particle swarm optimization, data mining.

1 Introduction

Data mining is a methodology for the extraction of new knowledge from data. In many applications of data mining technology, applying association rules are the most broadly discussed method. Early research in this field (such as Apriori) concentrated on Boolean association rules. Since real-world applications usually consist of quantitative values, the extraction of quantitative association rules have been regarded meaningful and crucial. Former quantitative association rule mining algorithm used discrete intervals to cover quantitative attribute. Nevertheless, intervals may not be concise and meaningful enough for human to obtain nontrivial knowledge. Since the comprehensibility of fuzzy rules by human users is a criterion in designing a fuzzy rule-based system, fuzzy rules with linguistic interpretation deals with "the boundary problem" naturally and can be introduced into data mining [1].

Some works have recently been done on the use of fuzzy sets in discovering association rules for quantitative attributes e.g. [2-5]. However, fuzzy sets usually determined by domain experts in existing approaches. This is not realistic because it is subjective and extremely hard for experts to specify fuzzy sets in a dynamic environment. In order to handle this problem, GA-based methods are employed in [6-8] to derive the fuzzy sets from given transactions. In [9], a 2-tuples linguistic representation model was introduced to fuzzy rules representation. On the concept of 2-tuples linguistic representation, [10] introduced a GA-based approach to 2-tuples fuzzy association rules.

* Corresponding author.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 621–630.
springerlink.com

It is well known that particle swarm optimization [11] is a population-based global optimization method based on a simple simulation of bird flocking or fish schooling behavior. The significant performances of PSO have been broadly discussed. Therefore, in this paper, we present a new fuzzy data mining algorithm for extracting both fuzzy association rules and MFs from quantitative transactions by means of particle swarm optimization. In this approach, the search space provided by the 2-tuples linguistic representation helps the particles to obtain appropriate MFs, which may extract more comprehensive fuzzy rules.

The rest of this paper is organized as follows. The concepts of 2-tuples fuzzy association rule are introduced in Section 2. In Section 3, we focus on the coding strategy of PSO and the definition of the fitness function. Section 4 depicts the algorithm of the proposed method. Finally, we consider some data from the United States census in year 2000 and conduct some experiments to test our approach.

2 Preliminaries: The 2-Tuples Fuzzy Association Rule

2.1 The Definition of Fuzzy Association Rule

A fuzzy association rule can be expressed as follows:

$$\text{IF } X=\{x_1, x_2, \ldots, x_p\} \text{ is } A=\{f_1, f_2, \ldots, f_p\},$$
$$\text{THEN } Y=\{y_1, y_2, \ldots, y_p\} \text{ is } B=\{g_1, g_2, \ldots, g_p\},$$

simplify as $A \Rightarrow B$, where X and Y are disjoint sets of attributes, i.e., $X \subset Attr$, $Y \subset Attr$, $X \cap Y = \varnothing$; f_i and g_i are fuzzy sets that relate to attribute x_i and y_i respectively. $Attr = \{a_1, a_2, \ldots, a_m\}$ denotes all attributes (items). The fuzzy support and confidence are given as follows:

Definition 1. [12] *Let $A_i = (a_{is}, f_{is})$ be an fuzzy item, where $a_{is} \in Attr$, f_i is a fuzzy set of a_{is} ,(i=1,2,…,k), then we call $P = A_1 \wedge A_2 \wedge \ldots \wedge A_k$ a fuzzy itemset.*

Definition 2. [12] *For any transaction t_i =(id, $< t_{i1}, t_{i2}, \ldots, t_{im} >$), the support of A_i in t_i can be defined as: $s(A_i, t_i) = \min\{\mu f_{is}(t_i) \mid A_i = (a_{is}, f_{is}), i = 1,2,\ldots,k\}$.* Where *$T = \{t_1, t_2, \ldots, t_n\}$ is a database of transactions, t_i represents the i th transaction.*

Then we call $\sigma(P,T) = \dfrac{\sum_{t_i \in T} s(P,t_i)}{|T|}$ the fuzzy support of P .

Definition 3. [12] *Define the fuzzy confidence of rule $A \Rightarrow B$ in database T be: $\varphi(A \Rightarrow B, T) = \dfrac{\sigma(A \wedge B, T)}{\sigma(A,T)}$, where A and B are both itemsets.*

Definition 4. [12] *Define i frequent itemset as* $L = \{A_1 \wedge A_2 \wedge ... \wedge A_i \mid \sigma(A_1 \wedge A_2 \wedge ... \wedge A_i, T) \geq \sigma_{\min}\}$, *where* $\sigma_{\min}$ *is the specified minimum fuzzy support.*

Explicitly, each frequent itemset L_i ($i \geq 2$) can be used to deriving association rules $(L_i - S) \Rightarrow S$, for each $S \subset L_i$ .Therefore, the problem of mining all fuzzy association rules converts into generating all frequent itemsets, which support and confidence are greater than user-specified.

2.2 The 2-Tuples Linguistic Representation

The 2-tuples linguistic representation scheme [9] introduces a new model for rule representation based on the concept of symbolic translation. The symbolic translation of a linguistic term is a number with in the interval [−0.5, 0.5) that expresses the domain of a linguistic term when it is moving between its two lateral linguistic term. Suppose a set of linguistic terms *U* representing a fuzzy partition. Formally, we have the pair as $(f_i,\ \alpha_i)$, where $f_i \in U,\ \alpha_i \in [-0.5,\ 0.5)$.

For example, consider a simple problem with five linguistic terms in the attribute "age" {Very young, Young, Middle, Old, Very old}. Fig. 1 depicts the five fuzzy sets in "age". Fig. 2 shows the symbolic translation of a linguistic term represented by the pair (*S*, −0.2). Moreover, we can use more comprehensive linguistic term such as"a bit smaller than Old"to replace (*S*, −0.2).

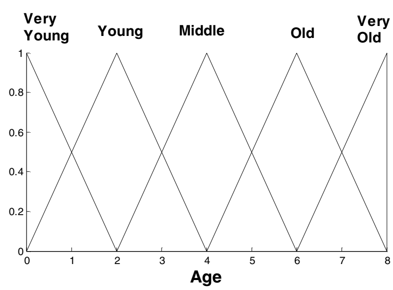


Fig. 1. Five triangle fuzzy sets in attribute "age"

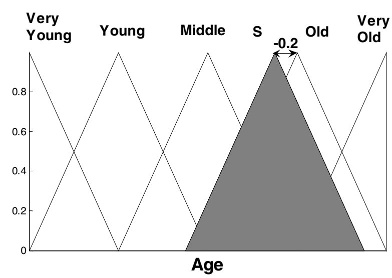


Fig. 2. An example of 2-tuples linguistic representation in "age"

Based on the concept, 2-tuples association rule in the following way:

If Age is (Middle, 0.3) then Weight is (High, −0.1).
In other words, the rule can be express as more comprehensive one:
If Age is (higher than Middle) then Weight is (a bit smaller than High).

3 NPSO-Based Method to Obtain the Membership Functions

3.1 Nonlinear Particle Swarm Optimization (NPSO)

As a population-based evolutionary algorithm, PSO [11] is initialized with a population of candidate solutions and the activities of the population are guided by some behavior rules. For example, let $X_i(t)= (x_{i1}(t), x_{i2}(t), \ldots, x_{iD}(t))$ ($x_{id}(t) \in [-x_{d\max}, x_{d\max}]$) be the location of the ith particle in the tth generation, where $x_{d\max}$ is the boundary of the dth search space for a given problem. The location of the best fitness achieved so far by the ith particle is denoted as $p_i(t)$ and the index of the global best fitness by the whole population, as $p_g(t)$. The velocity of ith particle is $V_i(t)= (v_{i1}(t), v_{i2}(t), \ldots, v_{iD}(t))$, where v_{id} is in $[-v_{d\max}, v_{d\max}]$ and $v_{d\max}$ is the maximal speed of dth dimension. The velocity and position update equations of the ith particle are given as follows:

$$v_{id}(t+1) = w \cdot v_{id}(t) + c_1 r_1 (p_{id} - x_{id}(t)) + c_2 r_2 (p_{gd} - x_{id}(t)) \tag{1}$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \tag{2}$$

where i=1,. . . , n and d=1,. . . , D. w, c_1, $c_2 \geq 0$. w is the inertia weight, c_1 and c_2 the acceleration coefficients, and r_1 and r_2 are randomly generated in the range [0, 1].

It is well known that a suitable value for the inertia weight provides a balance between the global and local exploration ability of the swarm. Base on the concept of decrease strategy [13, 14], a nonlinear inertia weight adaptation strategy was proposed in [15], which chooses lower value of w during the early iterations and maintains higher value of w than linear model. Experiment results demonstrated that nonlinear strategy enables particles to search the solution space more aggressively to look for "better areas", thus will avoid local optimum effectively.

The proposed adaptation of $w(t)$ is given as follows:

$$w(t) = \begin{cases} (1-\dfrac{2t}{iter_{\max}})^r \dfrac{(w_{initial}+w_{final})}{2} + \dfrac{(w_{initial}-w_{final})}{2}, t \leq \dfrac{iter_{\max}}{2} \\ (1-\dfrac{2(t-\frac{iter_{\max}}{2})}{iter_{\max}})^{\frac{1}{r}} \dfrac{(w_{initial}-w_{final})}{2} + w_{final}, \quad t > \dfrac{iter_{\max}}{2} \end{cases} \tag{3}$$

where $iter_{\max}$ is the maximum number of iterations, t the iteration generation and r>1 is the nonlinear modulation index.

Fig.3 shows the inertia weight variations with iterations for different values of r. In [15], we proved that a choice of r within [2-3] is normally satisfactory.

3.2 MFs Codification

In this paper, a real coding scheme is considered. We used membership functions in triangular shape because it is in general the most appropriate shape and the most widely used in fuzzy systems. Since the main task is to determine the appropriate setting of every fuzzy sets, each particle is a vector of real numbers with size n *m

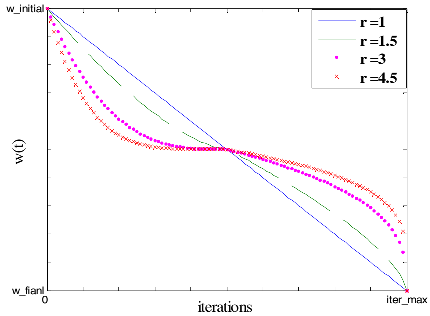


Fig. 3. Nonlinear model of inertia weight

(n items with m linguistic terms per item). As a result, every particle has the following form:

$(c_{11}, \ldots, c_{1m}, c_{21}, \ldots, c_{2m}, \ldots, c_{n1}, \ldots, c_{nm})$, where $c_{ij} \in [-0.5, 0.5)$, i=1, 2, …n, j=1, 2, …, m.

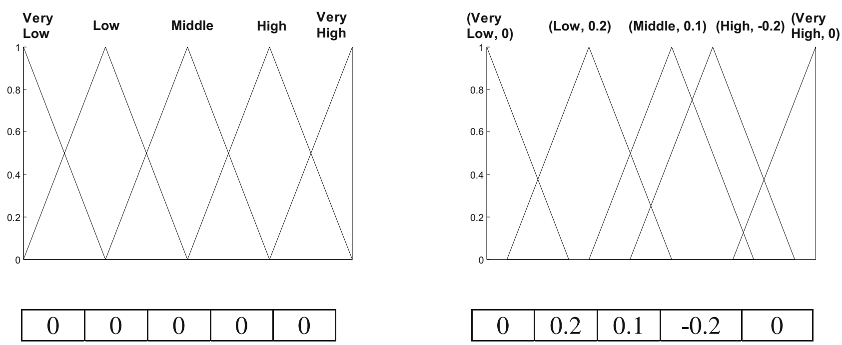


Fig. 4. Example of coding scheme

Fig 4 depicts an example of correspondence between a particle and it's associated MFs. Obviously, each particle is related to a series of fuzzy sets of itemsets. In order to include expert knowledge, the initial pool contains a particle that having all dimensions with value'0.0', and the other particles generated randomly in [−0.5, 0.5).

3.3 Fitness Evaluation

To evaluate a determined particle we will use the fitness functions defined in [10]:

$$f(x) = \frac{\sum_{P \in L_1} \sigma(P,T)}{suitability(x)} \tag{4}$$

where *suitability*(x) represents the shape suitability of the MFs from x, The suitability of the set of MFs in x is defined as:

$$suitability(x) = \sum_{k=1}^{n} (\sum_{i=1}^{m} \sum_{j=1}^{m} [\max(2\frac{overlap(R_{ki}, R_{kj})}{spanR_{ki}},\ 1)\ -1]+1) \tag{5}$$

where R_{kj} is the ith fuzzy set in the kth attribute, the $overlap(R_{ki}, R_{kj})$ is the overlap length of R_{ki} and R_{kj}, $spanR_{ki}$ is the span of R_{ki} and m denote the number of MFs for the ith attribute.

From the definition of the fitness function, the overlap factor represents the overlap ratio of the MFs for an item and the suitability factor can reduce the occurrence of redundant or separate shape of MFs. In other words, particles with higher fitness value can generate more appropriate fuzzy sets than lower ones.

4 Algorithm

Step1. Input training data, randomly initialize n particles and randomize the positions and velocities for entire population. Record the global best location p_g of the population and the local best locations p_i of the ith particle.;

Step2. Evaluate the fitness value of the ith particle through Eq.(4). If $(f(x_i)) < (f(p_i))$, set p_i=x_i as the so far best position of the ith particle. If $(f(x_i)) < (f(p_g))$, set p_g=x_i as the so far best position of the population;

Step3. Calculate the inertia weight through Eq.(3). Update the position and velocity of particles according to Eq.(1)and Eq.(2) (i=1,2, …, n);

Step4. Repeat Step2 and Step3 until Max number of generation or best solution;

Step5. Construct membership functions according to the best particle, extract all frequent itemsets and generate fuzzy association rules according to the algorithm proposed in [11].

5 Experiment Results

In this paper, the experiments conducted on 5,000 records from the United States census in the year 2000. We select 10 quantitative attributes from the database and five fuzzy sets have been defined for each attribute. As a matter of convenience, the proposed method, the GA approach [10], the conventional algorithm [11] are denoted as method ‘A’, ‘B’ , ‘C’, respectively.

5.1 Parameter Settings

(1) In the proposed method, $c_1 = c_2 = 2$; the population of particles is set as 40; $w_{initial}$=0.95, w_{final}=0.4, r=2.5.
(2) In method B, the population of GA is 50. The crossover and mutation probability is decreased linearly, where $p_{c\max}$ =0.8, $p_{c\min}$ =0.05, $p_{m\max}$ =0.5, $p_{m\min}$ =0.005.
(3) The number of iterations (PSO and GA) is set as 2000 and algorithms are implemented for 100 runs.

5.2 Results and Discussions

The analyses are taken with four statistical parameters (average fitness values, suitability, the number of large 1-itemsets and the number of interesting rules) for each method. Table 1 presents the results obtained by three different methods, where Fit for the fitness value, L_1 for the sum of the fuzzy support of the large 1-itemsets and Suit for the suitability. From the results listed in Table 1, it is observed that the proposed method achieved the best fitness value of searching over different minimum support. Fig 5 shows the average fitness values along with different numbers of evaluations for Method A and B. The number of large 1-itemsets point out that the proposed method is better than method B.

Table 1. Results obtained by three different strategies

Sup	Proposed approach			Method B			Method C		
	Fit	Suit	L_1	Fit	Suit	L_1	Fit	Suit	L_1
0.2	0.96	11.3	19	0.76	17.9	19	0.89	10.0	18
0.5	0.68	12.5	13	0.40	16.1	8	0.54	10.0	6
0.7	0.51	12.9	9	0.23	15.3	6	0.16	10.0	2
0.9	0.19	10.0	2	0.08	14.2	1	0.00	10.0	0

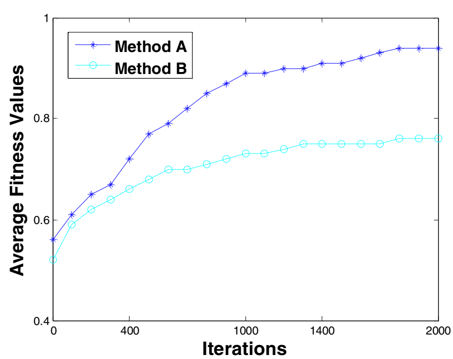


Fig. 5. The average fitness values along with different numbers of evaluations

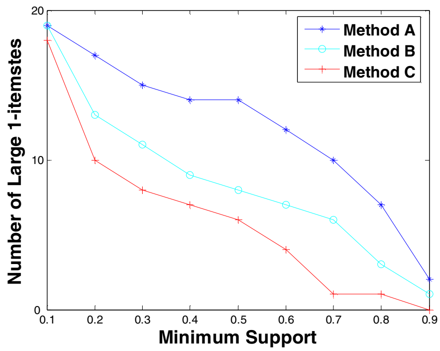


Fig. 6. Numbers of large 1-itemsets obtained for different minimum support

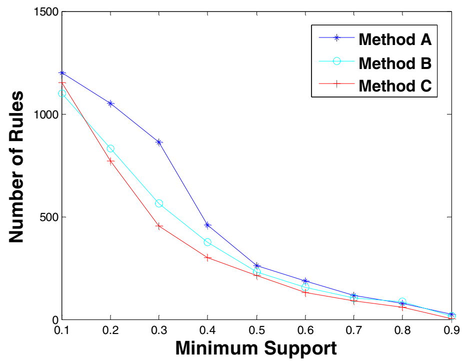


Fig. 7. Numbers of interesting rules obtained for different minimum support

Fig 6 and Fig 7 depict the number of large 1-itemsets and fuzzy association rules obtained by the different approaches. From these figures we can highlight that the proposed approach extracts the best number of fuzzy association rules for every minimum support. Moreover, although the derived number of fuzzy association rules decreased along with the increase of the minimum confidence value, the proposed method extracts about twice as fuzzy association rules as other approaches with all the values of the minimum support.

Some of the determined interesting fuzzy association rules are shown as below:

IF Age of person is (a bit higher than Old) AND education degree is (a bit smaller than High) THEN the number of persons in family is very low.

IF annual income of person is (a bit smaller than Very High) AND educational level is low THEN marital status is divorced.

6 Conclusion

In this paper, a new method is proposed to extract fuzzy association rules by using NPSO. The main task of PSO is to construct the appropriate fuzzy sets that cover the domain of quantitative attribute. The experiment results show that the proposed approach produces meaningful results and has reasonable efficiency. That is, NPSO-based approach can extract more interesting fuzzy association rules and large 1-itemsets than former methods.

Acknowledgements. The project is supported by the National Natural Science Foundation of Fujian Province, PRC (2008J0024); The key project of education department of Fujian province, PRC(JK2009017); The key project of institutes serving for the economic zone on the western coast of the Taiwan strait (2009(8)); The National Natural Science Foundation of Xiamen City, PRC (3502Z2093018).

References

1. Yan, P., Chen, G.-Q.: Mining non-redundant fuzzy quantitative association rules. In: Proc. in Fuzzy Information Processing Theories and Applications, pp. 391–396 (2003)
2. Hong, T., Kuo, C., Chi, S.: Trade-off between time complexity and number of rules for fuzzy mining from quantitative data. Internat. J. Uncertain Fuzziness Knowledge-Based Systems 9(5), 587–604 (2001)
3. Hong, T., Lee, Y.: An overview of mining fuzzy association rules. Studies in Fuzziness and Soft Computing 220, 397–410 (2008)
4. Kuok, C., Fu, A., Wong, M.: Mining fuzzy association rules in databases. SIGMOD Record 17(1), 41–46 (1998)
5. Lee, Y., Hong, T., Lin, W.: Mining fuzzy association rules with multiple minimum supports using maximum constraints. In: Negoita, M.G., Howlett, R.J., Jain, L.C. (eds.) KES 2004. LNCS (LNAI), vol. 3214, pp. 1283–1290. Springer, Heidelberg (2004)
6. Kaya, M., Alhajj, R.: Genetic algorithm based framework for mining fuzzy associations rules. Fuzzy Sets and Systems 152, 587–601 (2005)
7. Cai, G.-R., Chen, S.-L., et al.: Mining fuzzy association rules base on adaptive genetic algorithm. Operation Research, Management Science and Fuzziness 7(1), 73–82 (2007)
8. Hong, T., Chen, C., Wu, Y., Lee, Y.: A GA-based fuzzy mining approach to achieve a trade-off between number of rules and suitability of membership functions. Soft Comput. 10(11), 1091–1101 (2006)
9. Herrera, F., Martinez, L.: A 2-tuple fuzzy linguistic representation model for computing with words. IEEE Trans. Fuzzy Systems 8(6), 746–752 (2000)
10. Alcalá-Fdez, J., Alca, R., et al.: Learning the membership function contexts for mining fuzzy association rules by using genetic algorithms. Fuzzy Sets and System 160, 905–921 (2009)
11. Kennedy, J., et al.: Particle Swarm Optimization. In: International Conference on Neural Networks, pp. 1942–1948 (1995)

12. Cheng, J.-h., et al.: An algorithm for mining fuzzy association rules. MiniI-Micro Systems 4, 270–274 (1999)
13. Chatterjeea, A., Siarry, P.: Nonlinear inertia weight variation for dynamic adaptation in particle swarm optimization. Computers & Operations research, 859–871 (2006)
14. Shi, Y., et al.: A modified particle swarm optimizer. In: Procs. ICEC 1998, pp. 69–73. IEEE, Los Alamitos (1998)
15. Cai, G.-R., Chen, S.-L., et al.: Study on the Nonlinear Strategy of inertia weight in Particle Swarm Optimization Algorithm. In: International Conference on Natural Computation, pp. 683–687 (2008)

On $(\bar{\in}, \bar{\in} \vee \bar{q})$-Fuzzy Filters of Residuated Lattices

Yi-quan Zhu[1,⋆] Jian-ming Zhan[2], and Young Bae Jun[3]

[1] Division of Basic Sci., Guangzhou Maritime Univ., Guangzhou 510725, China
e-mail: zhuyiquan@126.com
[2] Dept of Math., Hubei Institute for Nationalities, Enshi 445000, Hubei, China
e-mail: zhanjianming@hotmail.com
[3] Dept of Math. Education, Gyeongsang National Univ., Chinju 660-701, Korea
e-mail: skywine@gmail.com

Abstract. The aim of this paper is to develop further the fuzzy filter theory of general residuated lattices. Mainly, we introduce the concept of $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filters in general residuated lattices, and derive some of their characterizations. Moreover, we discuss some relations between $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filters and several other special $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy filters.

Keywords: Residuated lattice, (Fuzzy) filter, $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy (implicative, positive implicative, fantastic and regular) filter.

1 Introduction

Residuated lattices, introduced by Dilworth and Ward in [1], are very basic algebraic structures among algebras associated with logical systems. In fact, many algebras have been proposed as the semantical systems of logical systems, for example, Boolean algebras, MV-algebras, BL-algebras, lattice implication algebras, MTL-algebras, NM-algebras and R_0-algebras, *etc.*, and they are all particular cases of residuated lattices. In addition, filter theory plays an important role in studying the interior structures of these algebras and the completeness of the corresponding logical systems. Therefore, it is meaningful to establish the filter theory of general residuated lattices for studying the common properties of the above mentioned algebras.

In [15], Y. Q. Zhu and Y. Xu extensively and profoundly discussed filters and fuzzy filters in general residuated lattices, including some special types of them. The aim of this paper is to develop further the fuzzy filter theory of general residuated lattices. We mainly introduce the concept of $(\bar{\in}, \bar{\in} \vee \bar{q}))$-fuzzy regular filters, which is a new type of the generalized fuzzy filters, and discuss some relative properties.

⋆ Corresponding author.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 631–640.
springerlink.com

2 Preliminaries

In this section, we recall some basic definitions and results which will be frequently used in the following and we shall not cite them every time they are used.

Definition 1 ([1,2]). *A residuated lattice is an algebraic structure* $L = (L, \wedge, \vee, \otimes, \rightarrow, 0, 1)$ *of type* $(2, 2, 2, 2, 0, 0)$ *satisfying the following axioms*:

(C1) $(L, \wedge, \vee, 0, 1)$ *is a bounded lattice.*
(C2) $(L, \otimes, 1)$ *is commutative semigroup* (*with the unit element* 1).
(C3) $(\otimes, \rightarrow)$ *is an adjoint pair.*

In a residuated lattice L, for any $x \in L$, we inductively define $x^1 = x, x^{k+1} = x^k \otimes x, k \in N$; $x' = x \rightarrow 0$, and $x'' = (x')'$, *etc.*

Proposition 1 ([2,10]). *In each residuated lattice L, the following properties hold for all $x, y, z \in L$:*

(P1) $(x \otimes y) \rightarrow z = x \rightarrow (y \rightarrow z)$.
(P2) $z \leq x \rightarrow y \Leftrightarrow z \otimes x \leq y$.
(P3) $x \leq y \Leftrightarrow z \otimes x \leq z \otimes y$.
(P4) $x \rightarrow (y \rightarrow z) = y \rightarrow (x \rightarrow z)$.
(P5) $x \leq y \Rightarrow z \rightarrow x \leq z \rightarrow y$.
(P6) $x \leq y \Rightarrow y \rightarrow z \leq x \rightarrow z, y' \leq x'$.
(P7) $y \rightarrow z \leq (x \rightarrow y) \rightarrow (x \rightarrow z)$.
(P8) $y \rightarrow x \leq (x \rightarrow z) \rightarrow (y \rightarrow z)$.
(P9) $1 \rightarrow x = x, x \rightarrow x = 1$.
(P10) $y \rightarrow x \leq x' \rightarrow y'$.
(P11) $x \leq y \Leftrightarrow x \rightarrow y = 1$.
(P12) $0' = 1, 1' = 0, x' = x''', x \leq x''$.

For the *regular residuated lattices*, MTL-*algebras*, BL-*algebras*, MV-*algebras* (*lattice implication algebras*) and NM-*algebras* (R_0-*algebras*) *etc.*, their definitions can be found in [2,11,15].

Definition 2 ([12]). *A fuzzy set F of the set X is a function $F : X \rightarrow [0, 1]$, and $U(F; t) := \{x \in X |\ F(x) \geq t\}$ is called a level subset of F for $t \in [0, 1]$.*

For the concepts of (fuzzy) filters, (fuzzy) implicative (positive implicative, fantastic and regular) filters and the relative results, readers may refer to [15].

Proposition 2 ([15]). *Let L be a residuated lattice, F a* (*fuzzy*) *subset of L. Then F is an* (*a fuzzy*) *implicative filter if and only if it is both a* (*fuzzy*) *positive implicative filter and a* (*fuzzy*) *fantastic/regular filter.*

We assume that the reader is acquainted with [7,8,15]; we shall refer to them all the time, and also shall use freely the concepts and terminologies appeared in [7,8,15].

In this paper, let L denote a residuated lattice unless otherwise specified.

3 $(\bar{\in}, \bar{\in} \vee \bar{q})$-Fuzzy Filters of Residuated Lattices

In this section, we enumerate some relative results on $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy filters obtained in some particular classes of residuated lattices, which still hold in general residuated lattices. Also, we will give some new results on $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy filters.

In [7,8], Ma and Zhan *et al.* introduced $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy filters and some special types in BL-algebras (R_0-algebras, respectively), obtained some useful results. We find that some of these results (including their proofs) are also the same available for general residuated lattices. Naturally, we could regard them as a part of the $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy filter theory of general residuated lattices. In order to need in the following studying, we only enumerate some relative results.

Let $x \in L$ and $t \in [0,1]$. A fuzzy set of L with the form

$$U(x;t) : L \to [0,1],\, U(x;t)(y) = \begin{cases} 1, & \text{if } y = x \\ 0, & \text{otherwise} \end{cases}$$

is said to be *a fuzzy point with support x and value t*.

A fuzzy point $U(x;t)$ is said to *belong to* (respectively, *be quasi-coincident with*) a fuzzy set F, written as $U(x;t) \in F$ (respectively, $U(x;t)qF$) if $F(x) \geq t$ (respectively, $F(x) + t > 1$). If $U(x;t) \in F$ or (respectively, and) $U(x;t)qF$, then we write $U(x;t) \in \vee qF$ (respectively, $U(x;t) \in \wedge qF$). The symbol $\overline{\in \vee q}$ means that $\in \vee q$ does not hold (see [7,8,9]).

Definition 3 ([7,8]). *A fuzzy set F of L is said to be an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy filter of L if for all $t, r \in [0,1]$ and $x, y \in L$,*
(G1) *$U(x \otimes y; \min\{t,r\})\overline{\in}F$ implies $U(x;t)\overline{\in} \vee \overline{q}F$ or $U(y;r)\overline{\in} \vee \overline{q}F$.*
(G2) *$U(y;r)\overline{\in}F$ implies $U(x;r)\overline{\in} \vee \overline{q}F$ with $x \leq y$.*

Theorem 1 ([7,8]). *The conditions* (G1) *and* (G2) *in* Definition 3*, respectively, are equivalent to the following conditions*:
(G3) $\max\{F(x \otimes y), 0.5\} \geq \min\{F(x), F(y)\}$*, $\forall x, y \in L$.*
(G4) $\max\{F(y), 0.5\} \geq F(x)$ *with $x \leq y$.*

Theorem 2. *A fuzzy set F of L is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy filter of L if and only if the following conditions are satisfied*:
(G5) $\max\{F(1), 0.5\} \geq F(x)$*, $\forall x \in L$.*
(G6) $\max\{F(y), 0.5\} \geq \min\{F(x), F(x \to y)\}$*, $\forall x, y \in L$.*

Proof. Necessity: assume that F is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-*fuzzy filter* of L and let $x, y \in L$. Then F satisfies the conditions (G3) and (G4) in Theorem 1. First, taking $y = 1$ in (G4), we have that the condition (G5) holds. Next, we show F satisfies (G6). In fact, since $x \leq (x \to y) \to y$, we have that $x \otimes (x \to y) \leq y$, and so it follows from (G4) that $\max\{F(y), 0.5\} \geq F(x \otimes (x \to y))$. Naturally, we have that $\max\{F(y), 0.5\} \geq 0.5$. Consequently, by (G3) we deduce that

$$\max\{F(y), 0.5\} \geq \max\{F(x \otimes (x \to y)), 0.5\} \geq \min\{F(x), F(x \to y)\},$$

which proves that (G6) holds.

Sufficiency: assume that F satisfies the conditions (G5) and (G6). First, we prove that F satisfies (G4). Let $\forall x, y \in L$ be such that $x \leq y$.

Then $x \to y = 1$, and so it follows from (G6) that $\max\{F(y), 0.5\} \geq \min\{F(x), F(x \to y)\} = \min\{F(x), F(1)\}$. Naturally, we have that $\max\{F(y), 0.5\} \geq 0.5$. Thus, by (G5) we see that

$$\begin{aligned} \max\{F(y), 0.5\} &\geq \max\{\min\{F(x), F(1)\}, 0.5\} \\ &\geq \min\{F(x), \max\{F(1), 0.5\}\} = F(x). \end{aligned} \tag{1}$$

This proves that F satisfies (G4). Now we show F satisfies (G3). For all $x, y \in L$, since $x \to (y \to (x \otimes y)) = (x \otimes y) \to (x \otimes y) = 1$, we have that $x \leq y \to (x \otimes y)$, and so it follows from (G4) that $\max\{F(y \to (x \otimes y)), 0.5\} \geq F(x)$. Naturally, we have that $\max\{F(x \otimes y), 0.5\} \geq 0.5$. On the other hand, by (G6) we have that $\max\{F(x \otimes y), 0.5\} \geq \min\{F(y), F(y \to (x \otimes y))\}$. Summing up the above results, we have that

$$\begin{aligned} \max\{F(x \otimes y), 0.5\} &\geq \max\{\min\{F(y), F(y \to (x \otimes y))\}, 0.5\} \\ &\geq \min\{F(y), \max\{F(y \to (x \otimes y)), 0.5\}\} \\ &\geq \min\{F(y), F(x)\}. \end{aligned} \tag{2}$$

This proves that F satisfies (G3). Hence, F is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy filter of L by Theorem 1.

Theorem 3. *A fuzzy set F of L is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy filter of L if and only if it satisfies the following condition:*

(GF) $x \leq y \to z \Rightarrow \max\{F(z), 0.5\} \geq \min\{F(x), F(y)\}$, $\forall x, y, z \in L$.

Proof. Suppose that F is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-*fuzzy filter* of L. Let $x, y, z \in L$ be such that $x \leq y \to z$. Then, by (G4) and (G6), we have $\max\{F(y \to z), 0.5\} \geq F(x)$ and $\max\{F(z), 0.5\} \geq \min\{F(y), F(y \to z)\}$. Obviously, $\max\{F(z), 0.5\} \geq 0.5$. Summing up the above results, we have that

$$\begin{aligned} \max\{F(z), 0.5\} &\geq \max\{\min\{F(y), F(y \to z)\}, 0.5\} \\ &\geq \min\{F(y), \max\{F(y \to z), 0.5\}\} \\ &\geq \min\{F(y), F(x)\}. \end{aligned} \tag{3}$$

This proves that F satisfies the condition (GF).

Conversely, suppose that F satisfies the condition (GF). Since $x \leq x \to 1$ for all $x \in L$, it follows from (GF) that $\max\{F(1), 0.5\} \geq \min\{F(x), F(x)\} = F(x)$. This shows that (G5) holds. On the other hand, since $x \to y \leq x \to y$ for all $x, y \in L$, it follows from (GF) that $\max\{F(y), 0.5\} \geq \min\{F(x \to y), F(x)\}$. This proves that (G6) holds. Hence F is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-*fuzzy filter* of L by Theorem 2.

Definition 4 ([7,8]). *Let F be an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy filter of L, and $x, y, z \in L$.*

(i) *F is called an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy implicative filter of L if it satisfies*

(G7) $\max\{F(x \to z), 0.5\} \geq \min\{F(x \to (z' \to y)), F(y \to z)\}$.

(ii) *F is called an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy positive implicative filter of L if it satisfies*

(G8) $\max\{F(x \to z), 0.5\} \geq \min\{F(x \to (y \to z)), F(x \to y)\}$.

(iii) *F is called an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy fantastic filter of L if it satisfies*

(G9) $\max\{F(((x \to y) \to y) \to x), 0.5\} \geq \min\{F(z \to (y \to x)), F(z)\}$.

Notation 1. In a residuated lattice L, every fuzzy (implicative, positive implicative, fantastic) filter F of L is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy (implicative, positive implicative, fantastic) filter, but the converses do not hold, respectively. The relevant examples can be found in [7,8]. However, if $F(x) \geq 0.5$ for all $x \in L$, then the converses hold, respectively.

Theorem 4 ([7,8]). *A fuzzy set F of L is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy (resp., implicative, positive implicative, fantastic) filter of L if and only if every non-empty level subset $U(F;t)$ is itself a (resp.,(an) implicative, positive implicative, fantastic) filter of L for all $t \in (0.5, 1]$.*

4 $(\bar{\in}, \bar{\in} \vee \bar{q})$-Fuzzy Regular Filters

In this section, we introduce the concept of $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filters of residuated lattices and investigate some of their properties.

Definition 5. *An $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy filter F of L is called an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filter of L, if it satisfies the following condition:*
(GFR) $\max\{F(x'' \to x), 0.5\} \geq F(1)$, $\forall x \in L$.

Notation 2. In general residuated lattices, $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filters exist, and also an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy filter may not be an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filter. The relevant examples can be found in [7,8].

Proposition 3 (Extension property). *Let F_1 and F_2 be $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy filters of L with $F_1 \leq F_2$ and $F_1(1) = F_2(1)$. If F_1 is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filter of L, then so is F_2.*

Proof. It is an immediate consequence of Definition 5.

Theorem 5. *In a residuated lattice L, the following assertions hold:*

(1) *Every fuzzy regular filter is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filter.*
(2) *If F is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filters with $F(1) \geq 0.5$, then F is indeed a fuzzy regular filter.*

Proof. It is an immediate consequence of Definition 5 and ([15], Definition 5.11).

Notation 3. The converse of Theorem 5(1) does not hold. In fact, an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filter may not be a fuzzy regular filter, even may not be a fuzzy filter. The relevant examples can be found in [7,8].

Now we describe some characterizations of $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filters in residuated lattices.

Theorem 6. *Let F be an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy filter of L. Then the following assertions are equivalent:*

(GR1) *F is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filter of L.*
(GR2) $\max\{F(y \to x), 0.5\} \geq F(x' \to y')$, $\forall x, y \in L$.
(GR3) $\max\{F(y' \to x), 0.5\} \geq F(x' \to y)$, $\forall x, y \in L$.

Proof. (GR1)⇒ (GR2). Suppose that F is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filter of L. Then F satisfies the condition (GFR). For all $x, y \in L$, using Proposition 1, we have that $x' \to y' \leq y'' \to x'' \leq y \to x''$, and so the following inequality holds:

$$x'' \to x \leq (y \to x'') \to (y \to x) \leq (x' \to y') \to (y \to x).$$

Thus, by (GF) in Theorem 3 we have that

$$\max\{F(y \to x), 0.5\} \geq \min\{F(x' \to y'), F(x'' \to x)\}.$$

This and $\max\{F(y \to x), 0.5\} \geq 0.5$ imply that

$$\begin{aligned} \max\{F(y \to x), 0.5\} &\geq \max\{\min\{F(x' \to y'), F(x'' \to x)\}, 0.5\} \\ &\geq \min\{F(x' \to y'), \max\{F(x'' \to x), 0.5\}\} \\ &\geq \min\{F(x' \to y'), F(1)\} \quad \text{(by (GRF))}. \end{aligned} \tag{4}$$

Similarly, we further obtain that

$$\begin{aligned} \max\{F(y \to x), 0.5\} &\geq \max\{\min\{F(x' \to y'), F(1)\}, 0.5\} \\ &\geq \min\{F(x' \to y'), \max\{F(1), 0.5\}\} \\ &= F(x' \to y') \quad \text{(by (G5))}. \end{aligned} \tag{5}$$

This shows that (GR2) holds.

(GR2)⇒ (GR1). Suppose that F satisfies the condition (GR2) and let $x \in L$. Since $x''' = x'$, we have that $x' \to (x'')' = x' \to x''' = 1$, and so $F(x' \to (x'')') = F(1)$. Thus, by the condition (GR2) we obtain that

$$\max\{F(x'' \to x), 0.5\} \geq F(x' \to x''') = F(1).$$

Therefore F is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filter of L by Definition 5.

(GR1)⇒ (GR3). Suppose that F is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filter of L and let $x, y \in L$. Using Proposition 1, we have that $x' \to y \leq y' \to x''$, and so we have the following inequality:

$$x'' \to x \leq (y' \to x'') \to (y' \to x) \leq (x' \to y) \to (y' \to x).$$

Thus, according to the methods of proving ((GR1)⇒ (GR2)), we can obtain that (GR3): $\max\{F(y' \to x), 0.5\} \geq F(x' \to y)$ for all $x, y \in L$.

(GR3)⇒ (GR1). Suppose that F satisfies the condition (GR3). Since $x' \to x' = 1$ for all $x \in L$, it following form (GR3) that

$$\max\{F(x'' \to x), 0.5\} \geq F(x' \to x') = F(1).$$

Hence F is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filter of L by Definition 5.

Theorem 7. *A fuzzy set F of L is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filter of L if and only if it satisfies*
(G5) $\max\{F(1), 0.5\} \geq F(x)$, $\forall x \in L$.
(GR4) $\max\{F(y' \to x), 0.5\} \geq \min\{F(z \to (x' \to y)), F(z)\}$, $\forall x, y, z \in L$.

Proof. Suppose that F is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filter of L. Then F is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy filter of L, and so the condition (G5) holds by Theorem 2. Now we show that F satisfies (GR4). In fact, by Theorem 2 we see that for any $x, y, z \in L$,

$$\max\{F(x' \to y), 0.5\} \geq \min\{F(z \to (x' \to y)), F(z)\}.$$

Next, by Theorem 6 we have that $\max\{F(y' \to x), 0.5\} \geq F(x' \to y)$, and obviously $\max\{F(y' \to x), 0.5\} \geq 0.5$. Consequently, we deduce that

$$\begin{aligned} \max\{F(y' \to x), 0.5\} &\geq \max\{F(x' \to y), 0.5\} \\ &\geq \min\{F(z \to (x' \to y)), F(z)\}. \end{aligned} \tag{6}$$

This proves that F satisfies (GR4).

Conversely, suppose that a fuzzy set F satisfies conditions (G5) and (GR4). First, taking $z = 1$ in (GR4), we have that

$$\max\{F(y' \to x), 0.5\} \geq \min\{F(x' \to y), F(1)\}.$$

This and $\max\{F(y' \to x), 0.5\} \geq 0.5$ deduce that

$$\begin{aligned} \max\{F(y' \to x), 0.5\} &\geq \max\{\min\{F(x' \to y), F(1)\}, 0.5\} \\ &\geq \min\{F(x' \to y), \max\{F(1), 0.5\}\} \\ &= F(x' \to y) \quad \text{(by (G5))}. \end{aligned} \tag{7}$$

Thus, we show that F satisfies (GR3). Next, we show F satisfies (G6). Let $x, y \in L$. Since $x \to y = x \to (1 \to y) = x \to (0' \to y)$, it follows from (GR4) that

$$\begin{aligned} \max\{F(y''), 0.5\} &= \max\{F(y' \to 0), 0.5\} \\ &\geq \min\{F(x \to (0' \to y)), F(x)\}. \end{aligned} \tag{8}$$

That is,

$$\max\{F(y''), 0.5\} \geq \min\{F(x \to y), F(x)\}. \tag{9}$$

Similarly, since $y'' = 1 \to (y' \to 0)$, it follows from (GR4) that

$$\max\{F(y), 0.5\} = \max\{F(1 \to y), 0.5\} \geq \min\{F(1 \to (y' \to 0)), F(1)\},$$

i.e.,

$$\max\{F(y), 0.5\} \geq \min\{F(y''), F(1)\}. \tag{10}$$

Obviously, we have that

$$\max\{F(y), 0.5\} \geq 0.5. \tag{11}$$

Thus, it follows from (10) and (11) that

$$\begin{aligned} \max\{F(y), 0.5\} &\geq \max\{\min\{F(y''), F(1)\}, 0.5\} \\ &\geq \min\{F(y''), \max\{F(1), 0.5\}\}. \end{aligned} \tag{12}$$

By (G5) and (11), we obtain that $\max\{F(y), 0.5\} \geq F(y'')$, and so that

$$\max\{F(y), 0.5\} \geq \max\{F(y''), 0.5\}. \tag{13}$$

Summing up the above (9) and (13), we prove that F satisfies condition (G6). Hence, F is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy filter of L by Theorem 2, and also it is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filter of L by Theorem 6 ((GR1)$\Leftrightarrow$ (GR3)).

Theorem 8. *A fuzzy set F of L is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filter of L if and only if it satisfies*
(G5) $\max\{F(1), 0.5\} \geq F(x)$, $\forall x \in L$.
(GR5) $\max\{F(y \to x), 0.5\} \geq \min\{F(z \to (x' \to y')), F(z)\}$, $\forall x, y, z \in L$.

Proof. It is similar to the proof of Theorem 7.

By using the level regular filters, we can characterize $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filters as follows:

Theorem 9. *A fuzzy set F of L is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filter of L if and only if every non-empty level subset $U(F;t)$ is a regular filter of L for any $t \in (0.5, 1]$.*

Proof. Let F be an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filter of L and $t \in (0.5, 1]$. Then $U(F;t)$ is a filter of L by Theorem 4. For all $x, y \in L$, if $x' \to y' \in U(F;t)$, then $F(x' \to y') \geq t$. According to Theorem 6 ((GR1)$\Rightarrow$ (GR2)), we have that $\max\{F(y \to x), 0.5\} \geq F(x' \to y') \geq t$. However, since $t > 0.5$, it follows that $F(y \to x) \geq t$, that is, $y \to x \in U(F;t)$. Thus, $U(F;t)$ is a regular filter of L by ([15], Theorem 5.3).

Conversely, if F is a fuzzy set of L such that $U(F;t)(\neq \emptyset)$ is a regular filter of L for all $t \in (0.5, 1]$, then F is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy filter of L by Theorem 4, and so that (G5) holds. In order to prove that F is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filter of L, now we need only to show that F satisfies (GR2) by Theorem 6. If not, then there exist $x, y \in L$ such that $\max\{F(y \to x), 0.5\} < F(x' \to y')$. Letting $t_0 = F(x' \to y')$, then we immediately see that $0.5 < t_0 \leq 1$ and $F(y \to x) < t_0$. This leads to $x' \to y' \in U(F;t_0)$ but $y \to x \notin U(F;t_0)$, so that $U(F;t_0)$ is not a regular filter of L by ([15], Theorem 5.3). However, this is a contradiction. Hence (GR2) holds, and F is indeed an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filter of L.

5 Relations among Some Generalized Fuzzy Filters

In this section, we discuss the relations among special types of $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy filters in general residuated lattices.

Ma and Zhan *et al*. established the relations among $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy implicative (positive implicative, fantastic) filters in BL-algebras and R_0-algebras, respectively. Some important results were respectively obtained in [7,8], in particular, it is proved that an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy filter is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy implicative filter if and only if it is both an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy positive implicative filter and an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy fantastic filter in BL-algebras. However, by using the characterizations on level filters of these special $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy filters, we can easily see that the same results also hold in general residuated lattices.

Theorem 10. *In a residuated lattice L, an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy filter of L is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy implicative filter if and only if it is both an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy positive implicative filter and an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy fantastic filter.*

Proof. It is a straightforward result of Theorem 4 and Proposition 2.

However, in a residuated lattice, an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy positive implicative filter or an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy fantastic filter may not be an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy implicative filter. The relevant examples can be found in [7,8].

In the following, we mainly discuss the relations between $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filters and the other special $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy filters in general residuated lattices. Our results are all similar to the cases of ordinary filters, and they can respectively be obtained from the corresponding results about ordinary filters, by using the characterizations on level subsets.

Theorem 11. *In residuated lattice L, every an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy fantastic filter is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filter.*

Proof. It is an immediate consequence of Theorem 4, Theorem 9 and ([15], Theorem 7.11).

Notation 4. In a residuated lattice L, an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filter may not be an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy fantastic filter. That is, the converse of Theorem 11 does not hold.

Similarly, we have the following:

Theorem 12. *In a* BL*-algebra L, $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy fantastic filters and $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filters are equivalent.*

Proof. It is an immediate consequence of Theorem 4, Theorem 9 and ([15], Theorem 7.12).

Theorem 13. *Let L be a residuated lattice, F a fuzzy set of L. Then F is an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy implicative filter if and only if it is both an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy positive implicative filter and an $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy regular filter.*

Proof. It is an immediate consequence of Theorem 4, Theorem 9 and Proposition 2.

Acknowledgements. This research is supported by the Natural Science Foundation of P. R. China (Grant No.60875034) and Research Foundation of Guangzhou Maritime University (Grant No.200912001).

References

1. Dilworth, R.P., Ward, M.: Residuated lattices. Trans. Amer. Math. Soc. 45, 335–354 (1939)
2. Hájek, P.: Metamathematics of fuzzy logic. Kluwer Academic Publishers, Dordrecht (1998)
3. Haveshki, M., Saeid, A.B., Eslami, E.: Some types of filters in BL-algebras. Soft Comput. 10, 657–664 (2006)
4. Jun, Y.B., Xu, Y., Zhang, X.H.: Fuzzy filters of MTL-algebras. Information Sciences 175, 120–138 (2005)
5. Kondo, M., Dudek, W.A.: Filter theory of BL-algebras. Soft Comput. 12, 419–423 (2008)
6. Liu, L.Z., Li, K.T.: Fuzzy Boolean and positive implicative filters of BL-algebras. Fuzzy Sets and Systems 152, 333–348 (2005)
7. Ma, X.L., Zhan, J.M., Dudek, W.A.: Some kinds of $(\bar{\in}, \bar{\in} \vee \bar{q})$-fuzzy filters of BL-algebras. Comput. and Math. with Appl. 58, 248–256 (2009)
8. Ma, X.L., Zhan, J.M., Jun, Y.B.: On $(\in, \in \vee q)$-fuzzy filters of R_0-algebras. Math. Log. Quart. 55, 452–467 (2009)
9. Pu, P.M., Liu, Y.M.: Fuzzy topology I: Neighourhood structure of a fuzzy point and Moore-Smith convergence. J. Math. Anal. Appl. 76, 571–599 (1980)
10. Turunen, E.: Boolean deductive systems of BL-algebras. Arch. Math. Logic 40, 467–473 (2001)
11. Xu, Y., Ruan, D., Qin, K.Y., Liu, J.: Lattice-valued logic. Springer, Heidelberg (2003)
12. Zadeh, L.A.: Fuzzy sets. Inform. and Control 8, 338–353 (1965)
13. Zhang, X.H.: On filters in MTL-algebras: Brief Report. Advances in Systems Science and Applications 7, 32–38 (2007)
14. Zhu, Y.Q., Roh, E.H., Zhang, Q.: On fuzzy implicative filters of implicative systems. Southeast Asian Bull. of Math. 27, 761–767 (2003)
15. Zhu, Y.Q., Xu, Y.: On filter theory of residuated lattices. Information Sciences (2010), doi:10.1016/j.ins.2010.05.034

On the Interpolation by Entire Function of the Exponential Type

Mei-ying Ren and Yong-lin Liu

Department of Mathematics and Computer,
Wuyi University, Wuyishan 354300, Fujian, China
e-mail: npmeiyingr@163.com

Abstract. By introducing the difference polynomial operator $P(\frac{1}{2h}\Delta_h)$, we obtain a kind of double-periodic entire $(0, P(\frac{1}{2h}\Delta_h))$ interpolation on equidistant nodes. we establish some equivalent conditions and give the explicit forms of some interpolation functions on the lacunary interpolation problem. The convergence of the interpolation operators is discussed.

Keywords: Difference polynomial operator, entire function, lacunary interpolation, convergence.

1 Introduction

Let $B^p_\sigma(1 \le p \le +\infty, \sigma > 0)$ denote the set of all entire functions f (on the complex plane C) of exponential type σ(i.e. $|f(z)| \le Me^{\sigma|y|}, z = x + iy \in C, M$ is a positive constant) which $f \in L^p(R)$ when restricted to R.

For a positive integer m, let $P(t) = \sum\limits_{j=0}^{m} c_j t^j$ $(c_j \in R, j = 0, 1, ..., m)$ is a real coefficient algebraic polynomial, $P(D)$ $(D = \frac{d}{dx})$ is a differential polynomial operator educed by $P(t)$.

For $f \in C_{2\pi},\ 0 < |h| < \frac{\pi}{2\sigma}$, let

$$\begin{aligned}
&\Delta^0_h f(x) = f(x),\\
&\Delta_h f(x) = \Delta^1_h f(x) = f(x+h) - f(x-h),\\
&\cdots\cdots\cdots\\
&\Delta^n_h f(x) = \Delta(\Delta^{n-1}_h f(x)) = \sum_{k=0}^{n} (-1)^k c^k_n f(x + (n-2k)h),\ (n \in Z, n \ge 2).
\end{aligned}$$

For the above-mentioned real coefficient algebraic polynomial $P(t)$, we call

$$P(\frac{1}{2h}\Delta_h) = \sum_{j=0}^{m} c_j \frac{1}{(2h)^j}\Delta^j_h$$

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 641–649.
springerlink.com

is a difference polynomial operator educed by $P(t)$. Naturally, if f has derivative of m orders then

$$\lim_{h\to 0^+} P(\frac{1}{2h}\Delta_h)f(x) = P(D)f(x).$$

The literature [1] studies entire $(0, P(D))$ interpolation problems. The literature [2] discusses so-called double-periodic entire $(0, m)$ interpolation problems and gets a similar conclusion as same as literature [1]. But,these interpolation problems all require the interpolated function with derivative of some orders. it can't be applied to the situation where in the node the interpolated function is not differentiable. Therefore its applicability is limited. So, the literature [3] uses high-order difference alternative to high-order derivative and studies trigonometric polynomial $(0, m)$ lacunary interpolation problems. Followed by the literature [4] studies the exponential type entire function of similar problems. The purpose of this paper is to follow the literature [1] what is used in Fourier analysis methods and Poisson summation formula, On the basis of the literature [2] ,we will discuss the so-called double-periodic entire $(0, P(\frac{1}{2h}\Delta_h))$ interpolation problems at equidistant nodes $x_{k,\sigma} = \frac{k\pi}{\sigma}$ $(k \in Z, \sigma > 0)$ through using the difference polynomial operator $P(\frac{1}{2h}\Delta_h)$ in place of m-order differential operator D^m.

2 Interpolation Problem and Its Main Conclusion

Our problems are:

(P_1) For any two given sets of complex $\{\alpha_k\}$ and $\{\beta_k\}$ $(k \in Z)$, which satisfying $\sum\limits_{k\in Z} |\alpha_k| < +\infty$, $\sum\limits_{k\in Z} |\beta_k| < +\infty$, whether or not there exists exponential type entire function $R(x) \in B^2_\sigma$ satisfying the condition:

$$R(x_{2k}) = \alpha_k, \quad (P(\frac{1}{2h}\Delta_h)R)(x_{2k+1}) = \beta_k, x_k = x_{k,\sigma} = \frac{k\pi}{\sigma}, k \in Z. \tag{1}$$

or equivalent to, whether or not there exists exponential type entire function $A(x), B(x) \in B^2_\sigma$, satisfying the condition:

$$\begin{cases} A(x_{2k}) = (P(\frac{1}{2h}\Delta_h)B)(x_{2k+1}) = \delta_{0,k}, \\ \quad B(x_{2k}) = (P(\frac{1}{2h}\Delta_h)A)(x_{2k+1}) = 0, \ k \in Z, \\ x_k = \frac{k\pi}{\sigma}, \quad k \in Z, \end{cases} \tag{2}$$

where $\delta_{0,0} = 1$, $\delta_{0,k} = 0$ as $k \in Z \setminus \{0\}$.

(P_2) If $f(x)$ is a bounded function on the real axis, setting $(R_\sigma f)(x) = \sum\limits_{k\in Z} f(x_{2k})A(x - x_{2k}) + \sum\limits_{k\in Z} \beta_k B(x - x_{2k})$, where $A(x), B(x) \in B^2_\sigma$ satisfying condition (2), whether the interpolation operator $(R_\sigma f)(x)$ is convergent.

In order to prove our main results, we need following lemmas.

Let $f \in L^1(R)$, if we denote the fourier transformation of the f by $\hat{f}$, then

$$\hat{f}(x) = \frac{1}{\sqrt{2\pi}} \int_R f(t)e^{-itx}dt, \ x \in R.$$

Lemma 1. *If $P(\frac{1}{2h}\Delta_h)$ is defined as above, $f(x) \in B^2_\sigma$, then*

$$(P(\frac{1}{2h}\Delta_h)f)^\wedge(t) = P(\frac{isinht}{h})f^\wedge(t).$$

It is easy to prove the lemma, omitted in this proof.

Lemma 2. *For $j = 1, 2, 3, \cdots$ and any real number t, we obtain*

$$\Delta_h^j costx = 2^j sin^j htcos(tx + \frac{j\pi}{2}), \ \Delta_h^j sintx = 2^j sin^j htsin(tx + \frac{j\pi}{2}).$$

Proof. It is easy to get the conclusion of the lemma by using mathematical induction.

Theorem 1. *(i) If $P(t)$ is a even real coefficient algebraic polynomial, $P(0) = 0$, and $P(\frac{isinht}{h}) \neq 0, \ t \in (0, \sigma]$, then $A(x), B(x) \in B^2_\sigma$ satisfy condition (2) if and only if $A(x), B(x)$ respectively have the following form :*

$$A(x) = \frac{2}{\sigma}\int_0^\sigma \frac{P((isinh(\sigma - t))/h)}{P((isinht)/h) + P((isinh(\sigma - t))/h)} costxdt, \tag{3}$$

$$B(x) = \frac{2}{\sigma}\int_0^\sigma \frac{cost(x - x_1)}{P((isinht)/h) + P((isinh(\sigma - t))/h)} dt, \tag{4}$$

Further, $R(x) \in B^2_\sigma$ satisfies the condition (1) if and only if

$$R(x) = \sum_{k\in Z} \alpha_k A(x - x_{2k}) + \sum_{k\in Z} \beta_k B(x - x_{2k}). \tag{5}$$

(ii)If $P(t)$ is a odd real coefficient algebraic polynomial, $P(\frac{isinht}{h}) \neq 0, \ t \in (0, \sigma]$, then in B^2_σ there don't exist entire functions $A(x)$ and $B(x)$ satisfying condition (2).

Proof. If $A(x)$ and $B(x)$ satisfy condition (2), let $A(x)$ and $(P(\frac{1}{2h}\Delta_h)A)(x)$ respectively replace $U(x)$ of the Theorem 2.2 which is in literature [1] and the Lemma 3 which is in literature [2], by Lemma 2.1, we get:

$$\begin{cases} A^\wedge(t) + A^\wedge(t + \sigma) = \frac{\sqrt{2\pi}}{\sigma}, & a.e. \ t \in (-\sigma, 0), \\ A^\wedge(t) + A^\wedge(t - \sigma) = \frac{\sqrt{2\pi}}{\sigma}, & a.e. \ t \in (0, \sigma), \\ P((isinht)/h)A^\wedge(t) - P((isinh(t + \sigma))/h)A^\wedge(t + \sigma) = 0, & a.e. \ t \in (-\sigma, 0), \\ P((isinht)/h)A^\wedge(t) - P((isinh(t - \sigma))/h)A^\wedge(t - \sigma) = 0, & a.e. \ t \in (0, \sigma). \end{cases} \tag{6}$$

Through solving the equation group (6), we have

$$A^\wedge(t) = \frac{\sqrt{2\pi}}{\sigma}\frac{P((isinh(\sigma-|t|))/h)}{P((isinht)/h)+P((isinh(\sigma-|t|))/h)}\ ,\ a.e.\ t\in(-\sigma,\sigma).$$

Since $A(x)\in B^2_\sigma,\ A(x)=\frac{1}{\sqrt{2\pi}}\int_R A^\wedge(t)e^{itx}dt$, so

$$A(x)=\frac{2}{\sigma}\int_0^\sigma \frac{P((isinh(\sigma-t))/h)}{P((isinht)/h)+P((isinh(\sigma-t))/h)}costxdt.$$

Let $B(x)$ and $(P(\frac{1}{2h}\Delta_h)B)(x)$ replace $V(x)$ of the Theorem 2.3 which is in literature [1] and $U(x)$ of the Lemma 2 which is in literature [2] respectively, according to Lemma 1, using as above similar method, we can get formula (4).

Contrary, if $A(x)$ is given by (3), then

$$A(x_{2k})=\frac{2}{\sigma}\int_0^\sigma \frac{P((isinh(\sigma-t))/h)}{P((isinht)/h)+P((isinh(\sigma-t))/h)}costx_{2k}dt. \tag{7}$$

Through appropriate variable substitution, we can get

$$A(x_{2k})=\frac{2}{\sigma}\int_0^\sigma \frac{P((isinht)/h)}{P((isinht)/h)+P((isinh(\sigma-t))/h)}costx_{2k}dt. \tag{8}$$

Through (7) and (8), we have $A(x_{2k})=\frac{1}{\sigma}\int_0^\sigma costx_{2k}dt=\delta_{0,k},\quad k\in Z.$

For $j,q,s\in N$, let

$$D_j(x):=(-1)^j\int_0^\sigma \frac{((isinh(\sigma-t))/h)^{2j}(sinht)^{2j}}{P((isinht)/h)+P((isinh(\sigma-t))/h)}costxdt,$$

$$E_{q,s}(x):=(-\frac{1}{h^2})^s\int_0^\sigma \frac{((isinh(\sigma-t))/h)^{2q}(sinht)^{2q+2s}}{P((isinht)/h)+P((isinh(\sigma-t))/h)}costxdt$$
$$+\int_0^\sigma \frac{((isinh(\sigma-t))/h)^{2q+2s}(sinht)^{2q}}{P((isinht)/h)+P((isinh(\sigma-t))/h)}costxdt.$$

By appropriate variable substitution, we have $D_j(x_{2k+1})=0$, $E_{q,s}(x_{2k+1})=0$.

By Lemma 2, we can get $P(\frac{1}{2h}\Delta_h)$ is a linear combination of various items which are shape such as $D_j(x)$ and $E_{q,s}(x)(j,q,s\in N)$. As $D_j(x_{2k+1})=0,\ E_{q,s}(x_{2k+1})=0,(j,q,s\in N)$, so $(P(\frac{1}{2h}\Delta_h)A)(x_{2k+1})=0,\ k\in Z.$

Similary, we have $B(x_{2k})=0,\ (P(\frac{1}{2h}\Delta_h)B)(x_{2k+1})=\delta_{0,k}\ ,\quad k\in Z.$

Let $R(x)\in B^2_\sigma$ satisfies the condition (1), we write $G(x)=R(x)-[\sum_{k\in Z}\alpha_k A(x-x_{2k})+\sum_{k\in Z}\beta_k B(x-x_{2k})]$, then $G(x_{2k})=0,\ (P(\frac{1}{2h}\Delta_h)G)(x_{2k+1})=0,\ k\in Z.$

As $G(x) \in B_\sigma^2$, so

$$G(x) = \frac{1}{\sqrt{2\pi}} \int_{-\sigma}^{\sigma} G^\wedge(t) e^{itx} dt, \ (P(\frac{1}{2h}\Delta_h)G)(x)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\sigma}^{\sigma} P(\frac{isinht}{h}) G^\wedge(t) e^{itx} dt, \ x \in R.$$

Let $G(x)$ and $(P(\frac{1}{2h}\Delta_h)G)(x)$ replace $V(x)$ of the Theorem 2.3 which is in literature [1] and $U(x)$ of the Lemma 3 which is in literature [2] respectively, according to Lemma 1, we get:

$$\begin{cases} G^\wedge(t) + G^\wedge(t+\sigma) = 0, \quad a.e.\ t \in (-\sigma, 0), \\ G^\wedge(t) + G^\wedge(t-\sigma) = 0, \quad a.e.\ t \in (0, \sigma), \\ P((isinht)/h)G^\wedge(t) - P((isinh(t+\sigma))/h)G^\wedge(t+\sigma) = 0, \quad a.e.\ t \in (-\sigma, 0), \\ P((isinht)/h)G^\wedge(t) - P((isinh(t-\sigma))/h)G^\wedge(t-\sigma) = 0, \quad a.e.\ t \in (0, \sigma). \end{cases} \tag{9}$$

Through solving the equation group (9), we have $G^\wedge(t) = 0, \ a.e.\ t \in (-\sigma, \sigma)$, so $G(x) = 0, \ x \in R$.

i.e. $R(x) = \sum\limits_{k \in Z} \alpha_k A(x - x_{2k)} + \sum\limits_{k \in Z} \beta_k B(x - x_{2k})$.

If $P(t)$ is a odd real coefficient algebraic polynomial, $P(\frac{isinht}{h}) \neq 0, \ t \in (0, \sigma]$, from as above proof, we know in B_σ^2 there don't exsit entire functions $A(x)$ and $B(x)$ satisfying condition (2). We now complete the proof of the theorem 1.

Lemma 3. ([5], [6]) (*Poisson Summation Formula*).

Let (i) $g \in L^1(R) \bigcap AC(R)$ or (ii) $g \in L^1(R) \bigcap C(R)$ with $g^\wedge \in L^1(R) \bigcap AC(R)$, then for any $\beta > 0$, we have

$$\sum_{k \in Z} g(\beta_k) = \frac{\sqrt{2\pi}}{\beta} \sum_{k \in Z} g^\wedge(\frac{2k\pi}{\beta}),$$

where $AC(R)$ is the set of all absolutely continuous function f on the R, $C(R)$ is the space of all uniformly continuous and bounded functions on the R.

Lemma 4. *If $P(t)$ satisfies the condition of the Theorem 2.3, $A(x)$ and $B(x)$ are given by formula (3) and (4) respectively, then $\|A\|_1 = O(\sigma^{-1}), \ \|B\|_r \leq c_2 \sigma^{-\frac{1}{r}} |P(\frac{isinh\sigma}{h})|^{-1}, \ \forall\ r > 1$, where constant $c_2 > 0$ is independ on σ.*

Proof. Setting

$$\varphi_\sigma(t) = \frac{P((isinh(\sigma - t))/h)}{P((isinht)/h) + P((isinh(\sigma - t))/h)},$$

then $\varphi_\sigma(\sigma) = 0, \ \varphi'_\sigma(\sigma) = 0$.

Integrating by parts twice , we get

$$A(x) = \frac{2}{\sigma}\int_0^\sigma \varphi_\sigma(t) costxdt = \frac{2}{\sigma}\int_0^\sigma \frac{1-costx}{x^2}\varphi''_\sigma(t)dt. \tag{10}$$

By a proper computing , we can find a constant $c > 0$, such that $|\varphi''_\sigma(t)| \leq c\ \sigma^{-2},\ (t > 0)$.

By appropriate variable substitution, we have,

$$\int_R \frac{1-costx}{tx^2}dx = \int_R \frac{1-cosu}{u^2}du < +\infty,\ (t > 0).$$

Hence, we obtain

$$\|A\|_1 = \int_R |A(x)|dx \leq 2\sigma^{-1}\int_0^\sigma tc_0c\sigma^{-2}dt = O(\sigma^{-1}),$$

where $c_0 = \int_R \frac{1-cosu}{u^2}du$.

Setting

$$\psi_\sigma(t) = \frac{1}{P((isinht)/h) + P((isinh(\sigma - t))/h)} - \frac{1}{P((isinh\sigma)/h)}.$$

by $P(0) = 0$, we have $\psi_\sigma(0) = \psi_\sigma(\sigma) = 0$.

Integrating by parts twice, we get

$$B(x + x_1) = \frac{2sin\sigma x}{\sigma x P((isinh\sigma)/h)} + \frac{2}{\sigma x^2}(cos\sigma x - 1)\psi'_\sigma(\sigma)$$

$$+\frac{2}{\sigma}\int_0^\sigma \frac{1-costx}{x^2}\psi''_\sigma(t)dt = B_1 + B_2 + B_3.$$

By a proper computing , we can find a constant $c_1 > 0$, such that

$$|\psi'_\sigma(\sigma)| \leq \frac{c_1\sigma^{-1}}{|P((isinh\sigma)/h)|},\quad |\psi''_\sigma(t)| \leq \frac{c_1\sigma^{-2}}{|P((isinh\sigma)/h)|}.$$

As $\int_R |\frac{sinu}{u}|^r du < +\infty$ ($\forall\ r > 1$), by Hölder-Minkowski inequality, we can get

$$\|B_1\|_r = O(\sigma^{-\frac{1}{r}}|P((isinh\sigma)/h)|^{-1}),\ \forall\ r > 1;$$

$$\|B_2\|_r = O(\sigma^{-\frac{1}{r}}|P((isinh\sigma)/h)|^{-1}),\ \forall\ r \geq 1;$$

$$\|B_3\|_r = O(\sigma^{-\frac{1}{r}}|P((isinh\sigma)/h)|^{-1}),\ \forall\ r \geq 1.$$

So, $\|B\|_r = O(\sigma^{-\frac{1}{r}}|P((isinh\sigma)/h)|^{-1})$.

Lemma 5. *Let $A(x)$ is given by formula (3). Then*
(*i*) $\sum_{k\in Z} A(x - x_{2k}) \equiv 1$; (*ii*) $\sum_{|x-x_{2k}|>\delta} |A(x - x_{2k})| \leq c_3(1 + \delta)\delta^{-2}\sigma^{-1}$, *where $c_3 > 0$ is a proper constant and δ is any positive number.*

Proof. (i) For a fixed $x \in R$, let $g_x(u) = A(x-u)$, $u \in R$, by literature [5], we have $A^\wedge(y) = 0$ as $|y| > \sigma$, and

$$[g_x]^\wedge(y) = \frac{1}{\sqrt{2\pi}} \int_R A(x-u)e^{-iyu}du$$

$$= \begin{cases} \frac{\sqrt{2\pi}}{\sigma} \frac{P((isinh(\sigma-|y|))/h)}{P((isinhy)/h)+P((isinh(\sigma-|y|))/h)} e^{-ixy}, & \text{if } |y| < \sigma, \\ 0, & \text{if } |y| \geq \sigma. \end{cases}$$

By Lemma 4 and $A(x) \in B_\sigma^2$, we have $g_x \in L^1(R) \bigcap AC(R)$, by Lemma 3, we have

$$\sum_{k\in Z} A(x - x_{2k}) = \sum_{k\in Z} g_x(\frac{2k\pi}{\sigma}) = \frac{\sigma}{\sqrt{2\pi}}[g_x]^\wedge(0) = 1.$$

(ii) For any $\delta > 0$, by formula (10), we get

$$A(x) = \frac{2}{\sigma}\int_0^\sigma \frac{1-costx}{x^2}\varphi_\sigma''(t)dt,$$

$$\sum_{|x-x_{2k}|>\delta} |A(x-x_{2k})| \leq (\frac{2}{\sigma}\int_0^\sigma |\varphi_\sigma''(t)|dt) \sum_{|x-x_{2k}|>\delta} |x-x_{2k}|^{-2}$$

$$= O(\sigma^{-2}) \sum_{|x-x_{2k}|>\delta} |x-x_{2k}|^{-2},$$

from

$$2\pi\sigma^{-1} \sum_{|x-x_{2k}|>\delta} |x-x_{2k}|^{-2} \leq 4\pi\delta^{-2}\sigma^{-1} + 2\int_0^{+\infty} u^{-2}du$$

$$= 4\pi\delta^{-2}\sigma^{-1} + 2\delta^{-1} \leq 2(1+\delta)\delta^{-2},\ (\sigma \geq 2\pi),$$

we have

$$\sum_{|x-x_{2k}|>\delta} |A(x-x_{2k})| \leq c_3(1+\delta)\delta^{-2}\sigma^{-1}.$$

Theorem 2. *Let $P(t)$ satisfies the conditions of Theorem 1. $f(x)$ is a bounded function on the real axis R, if $(R_\sigma f)(x) = \sum_{k\in Z} f(x_{2k})A(x-x_{2k}) + \sum_{k\in Z} \beta_k B(x-x_{2k})$, where $A(x)$ and $B(x)$ are given by formula (3) and (4) respectively, and $(\sum_{k\in Z} |\beta_k|^s)^{\frac{1}{s}} = o(\sigma^{\frac{1}{r}-1}|P(\frac{isinh\sigma}{h})|)$, for some real number $s > 1$, $\frac{1}{r} + \frac{1}{s} = 1$, then $(R_\sigma f)(x)$ converges to $f(x)$ at each continuity point x of $f(x)$ as $\sigma \to \infty$. Further, if $f(x)$ is a bounded and uniformly continuous function on real axis R, then $(R_\sigma f)(x)$ converges uniformly to $f(x)$.*

Proof. By Lemma 4, we have $A(x) \in B_\sigma^1$, by the nature of the entire function of exponential type (see literature [5]), we get

$$\sum_{k\in Z} |A(x-x_{2k})| \leq (1+\sigma)\|A\|_1 \leq (1+\sigma)\sigma^{-1}c_4 \leq 2c_4,\ (\sigma \geq 1, c_4 > 0$$

is a constant),

$$(\sum_{k\in Z} |B(x-x_{2k})|^r)^{\frac{1}{r}} \leq (1+\sigma)\|B\|_r \leq 2\sigma\|B\|_r \leq 2c_2\sigma^{1-\frac{1}{r}}|P((isinh\sigma)/h)|^{-1},\ \sigma \geq 1.$$

Let x is a point of continuity of $f(x)$, then for any $\varepsilon > 0$, there exists $\delta = \delta(x,\varepsilon) > 0$, such that $|f(x) - f(t)| < \varepsilon$ when $|t-x| \leq \delta$, then

$$|f(x) - (R_\sigma f)(x)| \leq \sum_{k\in Z} |f(x) - f(x_{2k})||A(x-x_{2k})| + \sum_{k\in Z} |\beta_k B(x-x_{2k})| \leq$$

$$(\sum_{|x-x_{2k}|\leq\delta} + \sum_{|x-x_{2k}|>\delta})|f(x) - f(x_{2k})||A(x-x_{2k})| + \sum_{k\in Z} |\beta_k B(x-x_{2k})| \leq$$

$$2c_4\varepsilon + O((1+\delta)\delta^{-2}\sigma^{-1}) + o(1).\ (\sigma \geq 2\pi).$$

At this we obtain the proof of the first part of the theorem. As to the second part of the theorem, as $f(x)$ is uniformly continuous on real axis R, then δ stated as above may be chosen such that it is independent of x. Thus, we can see that $(R_\sigma f)(x)$ converges uniformly to $f(x)$ as $\sigma \to \infty$.

Acknowledgements. This work described here is partially supported by the Science and Technology Foundation of Fujian Education Department (Grant no. JA09242).

References

1. Yong-ping, L.: On the trigonometric interpolation and the entire interpolation. Approx.Theory and Appl. 6(4), 85–106 (1990)
2. Xiang-qian, H.: On a 2-Periodic Lacunary Entrie Interpolation Problem. J. Math. Res. Exposition 15(3), 381–387 (1995)
3. Xie-hua, S.: A generalization of $(0, M)$ interpolation. J. Math. Res. Exposition 19(1), 9–17 (1999)
4. Mei-ying, W., Yong-ping, L.: Convergence rates for lacunary trigonometric interpoation in $L^p_{2\pi}$ spaces. Approx.Theory and Appl. 15(2), 90–99 (1999)
5. Butzer, P.L., Stens, R.L.: The Poisson summation formula, Whittake's cardial seriers and approximate integration. In: Canadian Math. Soci. Conference Proceedings, vol. 3, pp. 19–36 (1983)
6. Butzer, P.L., Nessel, R.J.: Fourier Analysis and Approximation. Academic Press, Basel (1971)
7. Sharma, A., Varma, A.K.: Trigonometric interpolation. Duck Math. J. 32, 341–357 (1965)

8. Gervais, R., Rahman, Q.I.: A extension of Carlson‘s theorem for entire functions of exponential type. J.Math.Analy.and Appl. 69, 585–602 (1979)
9. Zhen-yu, Z., Xiang-qian, H.: Approximation by $(0, m_1, m_2, \ldots, m_q)$ entrie interpolation operators in Besov space. Math. Appl. J. 16(2), 50–53 (2003)
10. Mei-ying, W.: A generalization of $(0, m)$-interpolation by entire functions. J. BJ Nor.Univ. 34(3), 313–318 (1998)
11. Sun, Y.-s.: Approximation Theory of Function. Normal University, Beijing (1989)
12. Ting-pan, X., Song-ping, Z.: Approximation of Real Function. Hangzhou University (1998)

A Multi-objective Algorithm Based on Discrete PSO for VLSI Partitioning Problem

Shao-jun Peng[1], Guo-long Chen[1,2,*], and Wen-zhong Guo[1]

[1] College of Mathematics and Computer Science,
Fuzhou University, Fuzhou, 350002, China
e-mail: psj35@126.com, {cgl,guowenzhong}@fzu.edu.cn
[2] Key Laboratory of Discrete Mathematics with Application of Ministry of Education,
Fuzhou, 350002, China
e-mail: mcgl@fzu.edu.cn

Abstract. The problem of circuit partitioning is a key phase in the physical design of VLSI. In this paper, we propose a multi-objective discrete PSO (DPSO) algorithm for the problem of VLSI partitioning. Moreover, a new strategy of heuristic local search is employed to accelerate the convergence. The main target of this multi-objective problem is optimizing the minimum cut and timing performance (delay) while area balance is taken as a constraint. The fitness function of phenotype sharing is used to evaluate solution by both pareto dominance and neighborhood density. The experimental results on ISCAS89 benchmarks are performed to validate the proposed algorithm. Compared with genetic algorithm (GA) and Tabu Search (TS) in literature [4], the proposed algorithm could obtain more markedly better solutions for bipartition problem.

Keywords: VLSI, Physical design, Partitioning, Multi-objective, Discrete PSO.

1 Introduction

Today, physical design is the most time consuming part in VLSI design flow, it also contains several stages. First of them is circuit partitioning whose results will be used directly in placement, routine and other stages of physical design. The most important objective of partitioning is to minimize the interconnections among the subdomains (cut-size). With current trends, partitioning with multi-objective which includes power, timing performance and area, in addition to minimum cut is in vogue. This work addresses the problem of VLSI circuit partitioning with the objective of minimizing delay and cut-size while considering the area balance constraint.

Circuit bipartition problem with the objective of minimum cut-size has been proved to be NP-hard [1]. For years many algorithms have been applied to solve multi-objective partitioning problem. According to the category of optimal method, these techniques are mainly classified into exact heuristic methods [2-3] and stochastic search methods which include TS algorithm [4], genetic algorithm [4-5], evolution algorithm [6-7] and so on. However, the shortages of low efficiency and

* Corresponding author.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 651–660.
springerlink.com

local optimum are universal existence on exact heuristic algorithms and most of the iterative improvement algorithms also have the problem of poor convergence accuracy and low evolution velocity [8].

As a swarm-based evolutionary method, PSO seems particularly suitable for multi-objective optimization mainly because of its high evolution velocity which the algorithm presents for single-objective [9]. Aiming at the aspects concerned above, we propose an effective algorithm based on DPSO to solve the multi-objective optimization in circuit bipartition. In order to overcome the shortage of DPSO algorithm that lacks the capacity of local search [10], an iterative refinement strategy based on fiduccia and mattheyses (FM) [11] is used to further improve cut-size objective of each particle. Moreover, local search for improving delay objective is also added while iterating. To decide the global best, a fitness value with phenotype sharing is defined, thus a non-dominance solution with lower neighborhood density would be selected [12].

The remainder of the paper is organized as follows. In the next section, the mathematical model of the multi-objective partitioning problem is formulated. Section III presents the proposed algorithm for partitioning problem in detail. Section IV summarizes the results obtained on the ISCAS89 benchmarks. Section V offers concluding remarks followed by acknowledgement.

2 Problem Description

Generally, a circuit can have multi-pin connections (nets) apart from two-pin and therefore it is better to represent it by a hypergraph $H(V,E)$, in which V={v_1 ,v_2 ,…,v_n} is a set of nodes (e.g., standard cells or gate) and E={e_1 ,e_2 ,…,e_m} is a set of hyperedges (e.g., nets). A net is a subset of nodes which are electrically connected by a signal. Considering the circuit which contains n nodes and m nets, the partitioning is to assign the n nodes to a specified number of subsets k (k=2 in this paper) satisfying prescribed properties.

Cut-size objective. *If subsets are denoted by V_1, V_2,…, V_k, the objective of cut-size can be written as follows:*

$$Minimize(\sum_{a=1}^{k}\sum_{b=a+1}^{k}\left|\{e \mid e\in E,(e\cap V_a\neq\phi)\bigcap(e\cap V_b\neq\phi)\}\right|)\cdot \tag{1}$$

This paper solves the problem of circuit bipartition.

Delay objective. The exact wire length can not estimate in circuit partitioning phase, therefore we used a simple path delay model in this paper [13]. It also can be extended much exactly if necessary.

In order to deal with a signal path, let P=(v_{p1} ,v_{p2} ,…,v_{pl}) be a directed path from v_{p1} to v_{pl}, if v_{p1}, v_{pl} is sequential nodes (or FF nodes) and v_{pj}($2\leq j\leq l-1$) is combinational nodes, P is a FF-FF path. If we assume that PI denotes the set of primary inputs, and PO denotes the set of primary outputs. For purposes of path timing analysis we treat the nodes of PI, PO, and FF as the end points of timing paths, i.e., the circuit delay is the longest combinational path (critical path) delay from any FF or PI output to any FF or PO input. We generically refer to timing paths as FF-FF paths [2]. A path-cut number of path P, denoted h(P), is the number of directed-edge cut which are included in the path P. The delay of a FF-FF path P can be written as follows:

$$d(p) = \sum_{v_i, v_j, v_k \in P} (d_{gate}(i) + d_{wire}(j,k)) + h(P) \times d_{inc} , \tag{2}$$

where $d_{gate}(i)$ denotes the delay of node v_i, $d_{wire}(j,k)$ denotes the delay of wire $e(v_j, v_k)$, d_{inc} denotes the delay increment of directed-edge cut. Then the objective of timing performance can be formally described as:

$$\text{Minimize } d_{critical\_path} = Max(d(P)), P \in H(V,E) \quad . \tag{3}$$

Area balance constraint. *This paper solves the problem of circuit bipartition and the area of all cells is assumed identical. Then the area balance constraint is given below:*

$$(1-\varphi) \times \frac{S}{2} \le S_i \le (1+\varphi) \times \frac{S}{2} \qquad (i = 1,2) , \tag{4}$$

where $\varphi (0 < \varphi < 1)$ denotes the area balance tolerance, S denotes the total area of all the cells in the circuit, S_i is the area of cells in partition V_i.

3 Proposed Algorithm

A multi-objective algorithm based on DPSO and local search strategy is now proposed for solving the circuit bipartition problem where the 2-tuple of objective defined above in Section 2, are minimized while considering the area balance constraint.

3.1 Encoding Scheme

This paper uses the scheme of 0-1 encoding for circuit bipartition problem. A particle is a feasible solution in circuit bipartition. Considering the circuit which contains n modules, a particle is a 0-1 array of n modules. For example, the particle i at time t can be represented by $X_i^t = (x_{i1}^t, x_{i2}^t, ..., x_{in}^t)$, where gene x_{ik}^t is 0 or 1. When x_{ik}^t value is 0, the corresponding module k is allocated to partition V_0, otherwise the corresponding module k is allocated to partition V_1.

3.2 Crossover and Mutation

Considering the problem with 0-1 encoding scheme, here we use the uniform crossover and random two-point exchange mutation operators.

Nowadays, there are many crossover operators for 0-1 encoding scheme. In order to have a complete and uniform search process, uniform crossover operator is used in this paper. DPSO is different from GA, there is only a new particle generated from crossover operator in DPSO for ensuring the population size remains unchanged. The number of genes c_{count} which are selected from guiding particle for offspring in crossover operation is determined by the product of total modules n and c_{ratio}. According to parameter c_{count}, randomly generating a 0-1 shield-array which satisfying the area balance constraint is the main step of this crossover operation. The c_{ratio} in this paper was set as 0.25.

Exchange mutation is a kind of mutation method, which is often used in many problems. Implementation of random two-point exchange mutation is described as: when two random exchange points are selected from a particle to be operated, an exchange between their corresponding modules will be made. In circuit bipartition problem, in order to avoid the invalid mutation, two random exchange modules must be located in different partitions of solution array.

3.3 DPSO Algorithm

This paper employs the new discrete position updating method [14] based on genetic operation for reference and proposes a DPSO algorithm for circuit bipartition problem. Position updating can be defined as follows:

$$X_i^t = w \oplus F_3(c_2 \oplus F_2(c_1 \oplus F_1(X_i^{t-1}, P_i^{t-1}), G_i^{t-1}), \tag{5}$$

where w is inertia weight, c_1 and c_2 are acceleration constants.

In the model, the position updating of particles consists of three sections thereinafter, where r_1, r_2, r_3 are random numbers on interval [0, 1) and F_1, F_2 denote the crossover operation of particles.

1. This section reflects the cognitive personal experience of particles:

$$\lambda_i^t = c_1 \oplus F_1(X_i^{t-1}, P_i^{t-1}) = \begin{cases} F_1(X_i^{t-1}, P_i^{t-1}), r_1 \le c_1 \\ X_i^{t-1}, r_1 > c_1 \end{cases}, \tag{6}$$

where c_1 denotes the crossover probability of particles and personal best solution.

2. This section reflects cooperative global experience of particles:

$$\delta_i^t = c_2 \oplus F_2(\lambda_i^t, G_i^{t-1}) = \begin{cases} F_2(\lambda_i^t, G_i^{t-1}), r_2 \le c_2 \\ \lambda_i^t, r_2 > c_2 \end{cases}, \tag{7}$$

where c_2 denotes the crossover probability of particles and global best solution.

3. This section reflects the velocity of particles:

$$X_i^t = w \oplus F_3(\delta_i^t) = \begin{cases} F_3(\delta_i^t), r_3 \le w \\ \delta_i^t, r_3 > w \end{cases}, \tag{8}$$

where F_3 denotes the mutation operation of particles, w denotes the mutation probability.

3.4 Local Search

Though the PSO has proved to be a global optimization algorithm, the DPSO used in this paper also lacks the capacity of local search [10]. In order to overcome the shortage concerned above, an iterative refinement strategy based on FM [11] is used to further improve cut-size objective of each particle. Moreover, local search for improving delay objective is also added while iterating. The local search process of particles consists of two sections thereinafter.

1. Local Search for reducing Cut-size of Particles

The approach which was presented by FM is an effective classical method for reducing cut-size in circuit partitioning problem. This section of local search employs the local migration strategy based on FM iterative heuristic for reference.

Definition 1. *As to the module i in circuit, The gain g(i) which is the number of nets by which the cut-size would decrease if module i was to be moved from its subset to another is defined as:*

$$g(i)=FS(i)-TE(i), \tag{9}$$

where FS(i) denotes the number of nets connected to module i and not connected to any other module in subset which contains module i, TE(i) denoted the number of nets that are connected to module i and not crossing the cut.

As to the particles which have updated the position by DPSO, the detail of the procedure is given bellow (L_1=0 is initialized, L_{1max} is the termination condition).

Step 1. Calculate the gain g(i) of each modules according to (9);

Step 2. Select the module k with a maximum gain g_{max} from all modules while considering the area balance constraint (4);

Step 3. If $g_{max}>0$, move the module k from its subset to another; $L_1=L_1+1$;

Step 4. If $L_1 \geq L_{1max}$ is satisfied, go to Step 5(next section of local search); otherwise, go to Step 1.

Considering the time complexity of the algorithm and the effect of local migration strategy for reducing cut-size, the termination condition L_{1max} in this paper was set as 2.

2. Local Search for reducing Delay of Particles

This part introduces the local search for improving delay objective.

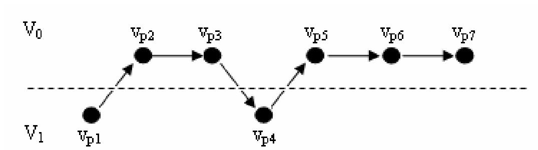


Fig. 1. Example of a FF-FF path P’

Definition 2. *As to the FF-FF path $P=(v_{p1}, v_{p2}, \ldots, v_{pl})$ in circuit, the $h(v_{pi})$ denotes the difference value of directed- edge cut in P by module v_{pi} was to be moved from its subset to another, the value range of $h(v_{pi})$ is {-2,-1,0,1,2}. For example, a FF-FF path P' in circuit is shown in figure 1. Then, the difference value $h(v_{pj})(1\leq j\leq 7)$ of all modules in P' are respectively {-1,0,0,-2,0,2,1}.*

As to the particles which have updated the position by local search for reducing cut-size, the detail of the procedure is given bellow.

Step 5. Calculate the difference value $h(v_{pi})$ of critical path modules;

Step 6. Select the module k with a minimum difference value h_{min} from all modules while considering the area balance constraint (4);

Step 7. If $h_{min}\leq 0$, move the module k from its subset to another.

3.5 Multi-objective Approach

To make a decision, a selection method is necessary, which should promote the swarm flying towards the real pareto front and distributing along the front as uniformly as possible. The fitness function of phenotype sharing defined in [12] is applied, in which a particle is evaluated by both pareto dominance and neighborhood density. For a global best position, a non-dominated solution with lower fitness value is selected. While iterating, a set A_1 contains non-dominated solutions with lowest fitness value in current swarm is maintained where a global best solution should be selected to affect the swarm flying and a set of non-dominated solutions A_2 is also maintained which is used to store non-dominated solutions generated in the whole search process. In particular, if several solutions have the same fitness value, we choose a random one. The detail of the fitness function is given below.

Definition 3. *As to particle x_i and x_j, the distance of the objective k is:*

$$f_k d_{ij} = \left| f_k(x_i) - f_k(x_j) \right| . \tag{10}$$

Therefore, the objective distance of particle x_i and x_j is given as:

$$fd_{ij} = f_1 d_{ij} + f_2 d_{ij} + \ldots + f_m d_{ij} , \tag{11}$$

where m is the dimension of the objective.

Definition 4. *The number D(i) which particle i is dominated is defined as follows:*

$$D(i) = \sum_{j=1}^{p} nd(i,j) , \tag{12}$$

where p is the population size, and nd(i,j) is 1 if particle i dominates particle j and 0 otherwise.

Definition 5. *The function of phenotype sharing is defined as follows:*

$$sh(fd_{ij}) = \begin{cases} 1, & if\ fd_{ij} \le \sigma_s, \\ 0, & otherwise \end{cases} \tag{13}$$

where σ_s denotes the sharing parameter.

Definition 6. *The neighbor density measure of particle i is defined as follows:*

$$N(i) = \sum_{j=1}^{p} sh(fd_{ij}), \tag{14}$$

where p denotes the population size.

Definition 7. *The fitness function of particle i is:*

$$F(i) = (1 + D(i))^{\alpha} \times (1 + N(i))^{\beta}, \tag{15}$$

where α and β are nonlinear parameters.

3.6 Algorithm Description

The details of algorithm are described as follows (the sets A_1, A_2 are defined in section 3.5):

Step 1: Load circuit net-list data and initialize the parameters of the algorithm;
Step 2: Randomly initialize population, calculate cut-size, delay and fitness value of each particle;
Step 3: Update personal best solution of each particle and the set A_1; randomly select a guide particle from A_1;
Step 4: Adjust the position and velocity of each particle according to (5)-(8);
Step 5: Update the position of each particle by local search mentioned in section 3.4;
Step 6: Calculate cut-size, delay and fitness value of each particle;
Step 7: Update the set A_2; if termination condition is satisfied, go to step 8; otherwise, go to step 3;
Step 8: Stop the algorithm and regard the set A_2 as the final solution.

4 Experimental Results

The experiments were run on a PC (1 CPU, 2.00GHz, 1.00GB RAM, Windows XP) and all the algorithms in experiments implemented in MATLAB. We tested the algorithms on the layouts of 4 sequential circuits in ISCAS89 benchmarks. The circuit parameters are summarized in Table 1. Generally, the FF-FF paths in sequential circuits are much more than that in combinational circuits. In order to save the storage space and run time of program, the FF-FF paths whose length is more than a certain threshold only be treated in this paper. The parameters of

Table 1. Specification of the benchmarks

Circuits	#Cells(including PI/PO)	#Nets	#FF	#Inverters	#FF-FF paths
s298	142	136	14	44	145
s820	331	312	5	33	492
s953	463	440	29	84	1156
s1238	554	540	18	80	3559

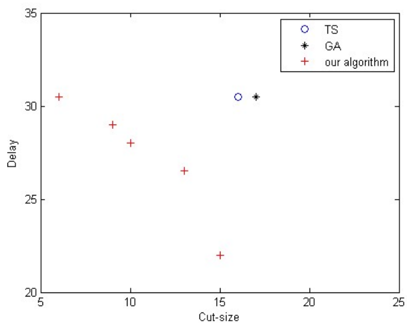


Fig. 2. Pareto front on circuit s298

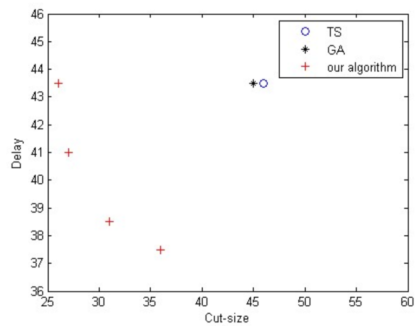


Fig. 3. Pareto front on circuit s820

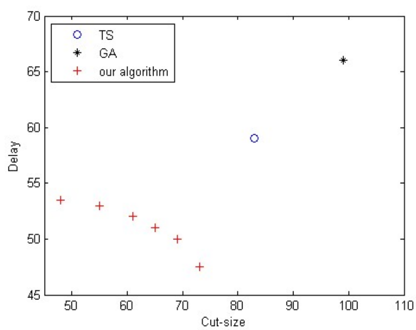


Fig. 4. Pareto front on circuit s953

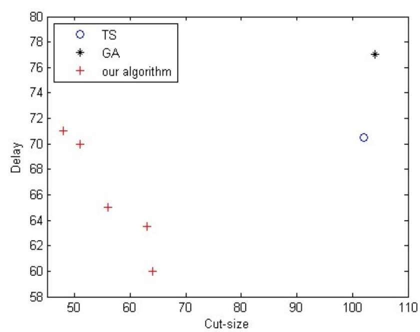


Fig. 5. Pareto front on circuit s1238

delay model were set as follows: d_{wire}=0.5, d_{inc}=10, for gate delay, if node v_i is inverter, $d_{gate}(i)$=1; if node v_i is PI, PO or FF, $d_{gate}(i)$=0; else $d_{gate}(i)$=2.

In order to validate the proposed algorithm, we compared it with the GA method [4] and the TS method [4] with area balance constraint ($\varphi = 0.2$). These two compared methods applied weighted sum approach to optimize the objectives, thus they obtained only one solution for a run. The parameters in the proposed algorithm were given as follows: population size p was 50, c_1 decreased linearly from 0.6 to 0.3, c_2 increased linearly from 0.2 to 0.4, w set to be 0.6, the maximum number of generations was 3000. The sharing parameters σ_s^1 and σ_s^2 were both set to be 2. The nonlinear parameters α and β were set to be 5 and 1, respectively. The best results of all three algorithms in 10 runs on each circuit are shown in Fig. 2-5.

From Fig. 2-5, as a direct representation, we can find that the pareto fronts of proposed algorithm in all 4 tested circuits dominated the GA and TS methods.Moreover, our algorithm is able to provide several feasible solutions for partitioning while other two algorithms only give out one.

To have a clearer view, we select three solutions uniformly from pareto fronts of each circuit. The exactly comparison of experimental results among three algorithms is shown in Table 2. H means the number of directed-edge cuts which are included in the critical path of solution. From the results, it is clear that the solutions of the proposed algorithm are all markedly better than that obtained from TS [4] and GA [4] methods among cut-size, delay and the directed-edge cut objectives.

Table 2. Comparison of experimental results(C=Cut-size, D=Delay, H=h($P_{critical\_path}$))

Circuit	Our algorithm			TS[4]			GA[4]		
	C	D	H	C	D	H	C	D	H
s298	6	30.5	1	16	30.5	1	17	30.5	1
	10	28	1						
	15	22	1						
s820	26	43.5	2	46	43.5	2	45	43.5	2
	27	41	2						
	36	37.5	2						
s953	48	53.5	3	83	59	4	99	66	5
	61	52	3						
	73	47.5	2						
s1238	48	71	3	102	70.5	3	104	77	3
	56	65	2						
	64	60	2						

5 Conclusion

In this paper, a multi-objective algorithm based on DPSO for VLSI circuit bipartition problem was proposed. In iterative process of DPSO, two heuristic local search strategies are used to further improve cut-size and delay objective of each particle, respectively. Experimental results on ISCAS89 benchmark circuits verified the feasibility and high-efficiency of the proposed algorithm by comparison with GA [4] and TS [4]. The future work will focus on multilevel partitioning problem by using the DPSO algorithm.

Acknowledgements. This work supported by the National Basic Research Program of China under Grant No. 2006CB805904, the National Natural Science Foundation of China under Grant No.10871221, Fujian Provincial Natural Science Foundation of China under Grant No.A0820002, Fujian Provincial Technology Innovation Platform Project of China under Grant No.2009J1007.

References

1. Wei, Y.C., Cheng, C.K.: Ratio cut partitioning for hierarchical designs. IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems 10, 911–921 (1991)
2. Kahng, A.-B., et al.: Local Unidirectional Bias for Cutsize-Delay Tradeoff in Performance-Driven Bipartitioning. IEEE Trans. on CAD of Integrated Circuits and Systems 23(4), 464–471 (2003)
3. Ababei, C., Navaratnasothie, S., Bazargan, K., Karypis, G.: Multi-objective circuit parti-tioning for cutsize and path-based delay minimization. In: Int. Conf. on CAD 2002, pp. 181–185 (2002)
4. Sait, S.M., El-Maleh, A.H., Al-Abaji, R.H.: General iterative heuristics for VLSI multiobjective partitioning. In: Int. Symp. on Circuits and Systems, vol. 5, pp. V497–V500 (2003)
5. Baruch, Z., Cret, O., Pusztai, K.: Genetic algorithm for circuit partitioning. In: Int. Conf. on Engineering of Modern Electric Systems: Computer Science and Control Systems, pp. 19–23 (1999)
6. Greiner, G.: Single and multi-objective evolutionary algorithms for graph bisectioning. In: 10th ACM SIGEVO workshop on Foundations of genetic algorithms, pp. 29–38 (2009)
7. Sait, S.M., El-Maleh, A.H., Al-Abaji, R.H.: Simulated evolution algorithm for multiobjective VLSI netlist Bi-partitioning. In: Int. Symp. on Circuits and Systems, vol. 5, pp. 457–460 (2003)
8. Peng, S.-j., Chen, G.-l., Guo, W.-z.: A Discrete PSO for Partitioning in VLSI Circuit. In: Int. Conf. on Computational Intelligence and Software Engineering, pp. 1–4 (2009)
9. Coello, C.A.C., Pulido, G.T., Lechuga, M.S.: Handling Multiple Objectives With Particle Swarm Optimization. IEEE Trans. on Evolutional Computing 8, 256–279 (2004)
10. Xu, J., Li, W., Wang, J.: Research on divergence analysis of discrete particle swarm optimization algorithm and its modification. Journal of System Simulation 21(15), 4676–4681 (2009)
11. Fiduccia, C.M., Mattheyses, B.M.: A linear-time heuristic for improving network partitions. In: 19th Design Automation Conference, pp. 175–181 (1982)
12. Guo, W., Chen, G., Huang, M., et al.: A discrete particle swarm optimization algorithm for the multiobjective permutation flowshop sequencing problem. In: The Second Int. Conf. on Fuzzy Information and Engineering, pp. 323–331 (2007)
13. Jie, H., Si-long, P.: Timing-driven multilevel partitioning algorithm with controlled path delay. Journal of CAD & Computer Graphics 20(5), 591–597 (2008)
14. Pan, Q.K., et al.: A discrete particle swarm optimization algorithm for the no-wait flow-shop scheduling problem. Computers & Operations Research 35, 2807–2839 (2008)

Generation and Recovery of Compressed Data and Redundant Data

Yu-ying Li[1,2], Li Zhang[2], and Kai-quan Shi[2]

Department of Computer and Information Engineering,
Ningde Normal University, Ningde 352100, China
e-mail: asszlyy@163.com
School of Mathematics and System Sciences,
Shandong University, Jinan 250100, China
e-mail: xitonglilun@126.com, shikq@sdu.edu.cn

Abstract. P-sets are a set pair which are composed of internal P-sets $X^{\bar{F}}$ and outer P-sets X^F . P-sets have dynamic characteristics. By using internal P-sets and deleted sets, the concepts of $\bar{F}$ -compressed data and $\bar{F}$ -redundant data are proposed. A data pair which is $((x)^{\bar{F}},(x)^{-})$ for data processing is proposed. Using it, some theorems are given such as the separation theorem, the generation theorems, recovery theorems about $\bar{F}$ -compressed data and $\bar{F}$ -redundant data. The measures about $\bar{F}$ -compressed degree, $\bar{F}$ -redundant degree and $\bar{F}$ -recovery degree are given. Finally, the application is provided about the system of typhoon pre-warning and search- rescue.

Keywords: P-sets, $\bar{F}$ -compressed data, $\bar{F}$ -redundant data, data generation, data recovery, $\bar{F}$ -data measure, application.

1 Introduction

In the computer data processing, data often requires to be compressed. Regardless of which method is used to compress data, the aim is to eliminate redundant data. Can we adopt a new mathematical model to understand the phenomenon of compressed data so as to obtain the theoretical characteristics and some new results? Can we apply this new mathematical theory and method to discuss the generation and recovery of compressed data and redundant data? In the literature which can be seen, no one gives these discussions.

In fact, some data elements $x_{p+1},x_{p+2},\cdots,x_q$ are deleted from data $(x)=\{x_1,x_2,\cdots,x_q\}$, then (x) changes into $(x)^{\bar{F}}=\{x_1,x_2,\cdots,x_p\}$, $p\le q$; this is equivalent to compressing data (x) to get data elements $x_1,x_2,\cdots,x_p$, thus compressed data $(x)^{\bar{F}}=\{x_1,x_2,\cdots,x_p\}$, $(x)^{\bar{F}}\subseteq(x)$ is generated. On the other hand, $x_{p+1},x_{p+2},\cdots,x_q$ are redundant data elements with respect to compressed data $(x)^{\bar{F}}$, they generate redundant data $(x)^{-}=\{x_{p+1},x_{p+2},\cdots,x_q\}$. $(x)^{\bar{F}}$ and $(x)^{-}$ compose $\bar{F}$ -data pair $((x)^{\bar{F}},(x)^{-})$. That (x) is compressed and changed into compressed data $(x)^{\bar{F}}$ is similar to the dynamic characteristics of P-sets (Packet sets)[1-2] created by Professor Kaiquan Shi in

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 661–671.
springerlink.com

2008. It is an interesting and important finding. In addition, there is no discussion and application on the deleted sets of X in the applied literatures of P-sets. Based on the theory of P-sets, the concepts of $\bar{F}$ -compressed data and $\bar{F}$ -redundant data are proposed in this paper. Using data pair $((x)^{\bar{F}},(x)^{-})$, the generation and recovery of $\bar{F}$ -compressed data and $\bar{F}$ -redundant data are discussed.

For the convenience of discussion, where P-sets structure, the dynamic characteristics and dependence are introduced into Section 2 for this paper as the theoretical preparation of discussion.

2 Prerequisites

Assumption 1. *X is a finite general set on U, U is a finite element universe, V is a finite attribute universe.*

Given a general set $X=\{x_1,x_2,\cdots,x_m\}\subset U$, and $\alpha=\{\alpha_1,\alpha_2,\cdots,\alpha_k\}\subset V$ is attribute set of X, $X^{\bar{F}}$ is called internal packet sets of X, called internal P-sets for short, moreover

$$X^{\bar{F}}=X-X^{-}, \tag{1}$$

X^{-} is called $\bar{F}$ -element deleted set of X, moreover

$$X^{-}=\{x\mid x\in X,\bar{f}(x)=u\,\bar{\in}\,X,\bar{f}\in\bar{F}\}, \tag{2}$$

if the attribute set α^{F} of $X^{\bar{F}}$ satisfies

$$\alpha^{F}=\alpha\cup\{\alpha'\mid f(\beta)=\alpha'\in\alpha,f\in F\}, \tag{3}$$

where $X^{\bar{F}}\neq\phi,\beta\in V,\beta\,\bar{\in}\,\alpha,f\in F$ turns β into $f(\beta)=\alpha'\in\alpha$.
Given a general set $X=\{x_1,x_2,\cdots,x_m\}\subset U$, and $\alpha=\{\alpha_1,\alpha_2,\cdots,\alpha_k\}\subset V$ is attribute set of X, X^{F} is called outer packet sets of X, called outer P-sets for short, moreover

$$X^{F}=X\cup X^{+}, \tag{4}$$

X^{+} is called F -element supplemented set, moreover

$$X^{+}=\{u\mid u\in U,u\,\bar{\in}\,X,f(u)\in x'\in X,f\in F\}, \tag{5}$$

if the attribute set $\alpha^{\bar{F}}$ of X^{F} satisfies

$$\alpha^{\bar{F}}=\alpha-\{\beta_i\mid\bar{f}(\alpha_i)=\beta_i\,\bar{\in}\,\alpha,\bar{f}\in\bar{F}\}, \tag{6}$$

where $\alpha^{\bar{F}}\neq\phi,\alpha_i\in\alpha,\bar{f}\in\bar{F}$ turns α_i into $\bar{f}(\alpha_i)=\beta_i\,\bar{\in}\,\alpha$.

The set pair which are composed of internal P-sets $X^{\bar{F}}$ and outer P-sets X^{F} are called P-sets (packet sets) generated by general set X, called P-sets for short, and it is expressed as the

$$(X^{\bar{F}},X^{F}), \tag{7}$$

where general set X is the ground set of $(X^{\bar{F}},X^{F})$.

It is should be pointed out here:

1°. The dynamic characteristic of formulas (1) - (6) is similar to the structure of $T = T + 1$ in the computer.

2°. $F = \{f_1, f_2, \cdots, f_m\}, \bar{F} = \{\bar{f}_1, \bar{f}_2, \cdots, \bar{f}_n\}$ are element transfer families [1-9, 10-12], $f \in F, \bar{f} \in \bar{F}$ are element transfers[1-9,10-12], $f \in F, \bar{f} \in \bar{F}$ are given functions.

3°. P-sets have dynamic characteristics. P-sets are set pair family composed by some set pairs. Moreover

$$\{(X_i^{\bar{F}}, X_j^F) \mid i \in \mathbf{I}, j \in \mathbf{J}\}, \tag{8}$$

where **I**, **J** are index sets, formula (8) is the representation of set pair family of P-sets.

4°. Internal P-sets have characteristic of one-directional dependence[3]. If α_1^F, α_2^F are attribute sets of $X_1^{\bar{F}}, X_2^{\bar{F}}$ respectively, moreover $\alpha_1^F \subseteq \alpha_2^F$, then there is $X_2^{\bar{F}} \subseteq X_1^{\bar{F}}$, $X_1^{\bar{F}}$ one-directionally depends on $X_2^{\bar{F}}$, moreover

$$X_2^{\bar{F}} \Rightarrow X_1^{\bar{F}}, \tag{9}$$

"$\Rightarrow$" in formula (9) comes from mathematical logic and reasoning, "$\Rightarrow$" is equivalent to "$\subseteq$".

3 The Generation-Recovery of $\bar{F}$ -Compressed Data and $\bar{F}$ -Redundant Data

Assumption 1. *In order to discuss easily, the finite general sets $X, X^{\bar{F}}, X^-$ in section 2 are expressed by (x), $(x)^{\bar{F}}$, $(x)^-$ respectively, or, $(x) = X$, $(x)^{\bar{F}} = X^{\bar{F}}$, $(x)^- = X^-$.*

Assumption 2. *The general form of data pair sequence is expressed by $((x)_i^{\bar{F}}, (x)_i^-)$. Specially, the initial states of $\bar{F}$ -compressed data and $\bar{F}$ -redundant data are $(x)_\circ^{\bar{F}} = (x)$, $(x)_\circ^- = \phi$.*

Definition 1. *$(x) = \{x_1, x_2, \cdots, x_q\} \subset U$ is called data on U, $x_i \in (x)$ is called data element of data (x), i=1,2,$\cdots$,q; if (x) has attribute set α, moreover*

$$\alpha = \{\alpha_1, \alpha_2, \cdots, \alpha_k\}. \tag{10}$$

Definition 2. *$(x)^{\bar{F}} = \{x_1, x_2, \cdots, x_p\} \subset U$ is called $\bar{F}$ -compressed data of data (x), moreover*

$$(x)^{\bar{F}} = (x) - (x)^-, \tag{11}$$

$(x)^- = \{x_{p+1}, x_{p+2}, \cdots, x_q\} \subset U$ is called $\bar{F}$ -redundant data of data (x), moreover

$$(x)^{-}=\{x \mid x \in (x), \bar{f}(x)=x' \overline{\in} (x), \bar{f} \in \bar{F}\}, \tag{12}$$

if the attribute set α^F *of* $(x)^{\bar{F}}$ *and attribute set* α *of* (x) *satisfy*

$$\alpha^F = \alpha \bigcup \{\alpha' \mid f(\beta)=\alpha' \in \alpha, f \in F\}, \tag{13}$$

where $\beta \in V, \beta \overline{\in} \alpha, p<q$ *;* $p,q \in \mathrm{N}^+$.

Definition 3. $((x)^{\bar{F}},(x)^{-})$ *is called* $\bar{F}$ *-compressed-redundant data pair of data* (x)*, if* $(x)^{\bar{F}}$ *is* $\bar{F}$ *-compressed data of data* (x)*,* $(x)^{-}$ *is* $\bar{F}$ *-redundant data of data* (x).

By formulas (1)-(3) and definitions 1-3, we get that:

Definition 4. $(x)_n^{\bar{F}}$ *is called the* n *level of* $\bar{F}$ *-compressed data of data* (x)*,* $(x)_n^{-}$ *is called the* n *level of* $\bar{F}$ *-redundant data of data* (x)*, if* $(x)_n^{\bar{F}}$ *and* $(x)_n^{-}$ *are obtained from attribute set* α *of* (x) *supplemented by attributes at the* n *time, satisfy*

$$\alpha \subseteq \alpha_1^F \subseteq \alpha_2^F \cdots \subseteq \alpha_{n-1}^F \subseteq \alpha_n^F . \tag{14}$$

Using definitions 1-4 and discussions in section 2, we get that:

Theorem 1. *(The separation theorem of* $\bar{F}$ *-compressed data and* $\bar{F}$ *-redundant data) If attribute set* α *of data* (x) *is supplemented by some attribute thus changes into* α^F *, and* $\alpha \subseteq \alpha^F$ *, then* $\bar{F}$ *-compressed data* $(x)^{\bar{F}}$ *and* $\bar{F}$ *-redundant data* $(x)^{-}$ *generated by data* (x) *satisfy*

$$(x)^{-}=(x)-(x)^{\bar{F}} . \tag{15}$$

The proof can be obtained from definition 2. It is omitted.

From formula (11) in definition 2 and formula (15) in theorem 1, we can get that:

Proposition 1. $\bar{F}$ *-compressed data* $(x)^{\bar{F}}$ *and* $\bar{F}$ *-redundant data* $(x)^{-}$ *of data* (x) *are separated and discovered to each other.*

Corollary 1. *If* $\bar{F}=\phi$*, then* $(x)^{\bar{F}}=(x)$ *; and vice versa.*
Where $\bar{F}=\{\bar{f}_1,\bar{f}_2,\cdots,\bar{f}_n\}$ is element transfer family, $\bar{f}_i \in \bar{F}$, $i \in \{1,2,\cdots,n\}$.

In fact, $\bar{F}=\phi$, then $(x)^{-}=\{x \mid x \in (x), \bar{f}(x)=x' \overline{\in} (x), \bar{f} \in \bar{F}\}=\phi$; Based on theorem 1, $(x)^{\bar{F}}=(x)-(x)^{-}=(x)$.

Corollary 2. *If* $(x)^{\bar{F}}$ *and* $(x)^{-}$ *are* $\bar{F}$ *-compressed data and* $\bar{F}$ *-redundant data of data* (x) *respectively, moreover*

$$(x)^{\bar{F}} \bigcap (x)^{-}=\phi . \tag{16}$$

The proof can be obtained from theorem 1. It is omitted.

By formula (16), we get:

Corollary 3. *If* $(x)^{\bar{F}}$ *and* $(x)^{-}$ *are* $\bar{F}$ *-compressed data and* $\bar{F}$ *-redundant data of data* (x) *respectively, then*

$$\mathbf{IDE}\{(x)^{\bar{F}},(x)^{-}\}, \quad (17)$$

where $\mathbf{IDE}$ =identification [1-2].

Theorem 2 *(Existence theorem of* $\bar{F}$ *-compressed data and* $\bar{F}$ *-redundant data). Suppose* α *is attribute set of data* (x)*, if there exists attribute set* α^{*}*, moreover* α^{*}*,* α *satisfy*

$$\alpha^{*}-\alpha\neq\phi, \quad (18)$$

then there exists data $(x)^{*}$ *and* $(x)^{\circ}$*,* $(x)^{*}$ *is* $\bar{F}$ *-compressed data of* (x)*,* $(x)^{\circ}$ *is* $\bar{F}$ *-redundant data of* (x)*.Where* $(x)^{*}$ *has attribute set* α^{*}*.*

Proof. As $\alpha^{*}-\alpha\neq\phi$, then $\alpha\subseteq\alpha^{*}$. Based on formulas (1)-(3), there is: being supplemented by some attributes, α changes into α^{*}. Because α , α^{*} are attribute sets of (x) and $(x)^{*}$ respectively, by definitions 1 and 2, there is: $(x)^{*}\subseteq(x)$, $(x)^{*}$ is data of (x) , $(x)^{*}=(x)^{\bar{F}}$. So $(x)^{*}$ exists. Suppose $(x)^{\circ}=(x)-(x)^{\bar{F}}$, then $(x)^{\bar{F}}=(x)-(x)^{\circ}$. By definition 2, then $(x)^{\circ}$ is $\bar{F}$ -redundant data of (x) , $(x)^{\circ}=(x)^{-}$. So $(x)^{\circ}$ exists.

Theorem 3 *(Generation theorem of* $\bar{F}$ *-compressed-redundant data pair sequence). If attribute set* α *of* (x) *is supplemented by attributes for n times, the attribute sets sequence of* α *obtained from it satisfy that:* $\alpha\subseteq\alpha_1^F\subseteq\alpha_2^F\cdots\subseteq\alpha_i^F\cdots\subseteq\alpha_{n-1}^F\subseteq\alpha_n^F$ *, then data* (x) *generates* $\bar{F}$ *-compressed-redundant data pair sequence* $((x)_i^{\bar{F}},(x)_i^{-})$ *; moreover* $(x)_i^{\bar{F}}$ *satisfies*

$$(x)_n^{\bar{F}}\subseteq(x)_{n-1}^{\bar{F}}\cdots\subseteq(x)_i^{\bar{F}}\cdots\subseteq(x)_2^{\bar{F}}\subseteq(x)_1^{\bar{F}}\subseteq(x), \quad (19)$$

where $i=1,2,\cdots,n$ *.*

Proof. When $i=1$, the attribute set α of (x) is supplemented by attributes to get attribute set α_1^F , $\alpha\subseteq\alpha_1^F$; By definitions 1-3, we get $\bar{F}$ -compressed data $(x)_1^{\bar{F}}$ and $\bar{F}$ -redundant data $(x)_1^{-}$ of (x) . Meanwhile $\bar{F}$ -compressed-redundant data pair $((x)_1^{\bar{F}},(x)_1^{-})$ is generated, moreover $(x)_1^{\bar{F}}\subseteq(x)$. Apparently, formula (19) is established. Suppose when $i=n-1$, the conclusion is established, moreover they satisfy $(x)_{n-1}^{\bar{F}}\subseteq(x)_{n-2}^{\bar{F}}\cdots\subseteq(x)_i^{\bar{F}}\cdots\subseteq(x)_2^{\bar{F}}\subseteq(x)_1^{\bar{F}}\subseteq(x)$ .When $i=n$, adding attributes to attribute set α_{n-1}^F of data $(x)_{n-1}^{\bar{F}}$, there is: α_n^F , $\alpha_{n-1}^F\subseteq\alpha_n^F$. Based on definitions 1-3, $(x)_n^{\bar{F}}$ and $(x)_n^{-}$ generated by data $(x)_{n-1}^{\bar{F}}$ are $\bar{F}$ -compressed data $(x)_n^{\bar{F}}$ and $\bar{F}$ -redundant data $(x)_n^{-}$ of $(x)_{n-1}^{\bar{F}}$ respectively, $((x)_n^{\bar{F}},(x)_n^{-})$ is $\bar{F}$ -compressed-redundant data pair of data $(x)_{n-1}^{\bar{F}}$, moreover $(x)_n^{\bar{F}}\subseteq(x)_{n-1}^{\bar{F}}$. Apparently, formula (19) is established. By the mathematical induction, the theorem can be proved.

Corollary 4. *If* $(x)_i^{\bar{F}}$ *is the* i *level of* $\bar{F}$ *-compressed data generated by data* (x), *then* $(x)_{i-1}^{\bar{F}}$ *one-directionally depends on* $(x)_i^{\bar{F}}$, *or*

$$(x)_i^{\bar{F}} \Rightarrow (x)_{i-1}^{\bar{F}}, \tag{20}$$

where $i = 1,2,\cdots,n$. *Specially, when* $i = 1$, *data* (x) *one-directionally depends on* $\bar{F}$ *-compressed data* $(x)^{\bar{F}}$, *or*

$$(x)^{\bar{F}} \Rightarrow (x), \tag{21}$$

where $(x) = (x)_{\circ}^{\bar{F}}$, $(x)^{\bar{F}} = (x)_1^{\bar{F}}$.
Corollary 4 can be obtained from formula (19) in theorem 3, one-direction dependence characteristic of P-sets and formula (9) directly, the proof is omitted.

Corollary 5. *If* $(x)_i^{\bar{F}}$ *is the* i *level of* $\bar{F}$ *-compressed data generated by data* (x), *then*

$$\mathbf{IDE}\{(x)_i^{\bar{F}},(x)_{i-1}^{\bar{F}}\}, \tag{22}$$

where $i = 1,2,\cdots,n$. *Specially, when* $i = 1$, *data* (x) *and* $\bar{F}$ *-compressed data* $(x)^{\bar{F}}$ *satisfy*

$$\mathbf{IDE}\{(x),(x)^{\bar{F}}\}, \tag{23}$$

where $(x) = (x)_{\circ}^{\bar{F}}$, $(x)^{\bar{F}} = (x)_1^{\bar{F}}$.
Corollary 5 can be directly obtained from formulas (19),(21) . The proof is omitted.

Theorem 4. *(Attribute recovery theorem of* $\bar{F}$ *-compressed data) The necessary and sufficient condition of* $\bar{F}$ *-compressed data* $(x)^{\bar{F}}$ *being restored to data* (x) *is that attribute set* α^F *of* $(x)^{\bar{F}}$ *and attribute set* α *of* (x) *satisfy*

$$\alpha^F - \{\beta_i \mid \alpha_i \in \alpha^F, \bar{f}(\alpha_i) = \beta_i \bar{\in} \alpha^F, \bar{f} \in \bar{F}\} = \alpha. \tag{24}$$

Proof. $\mathbf{1}^{\circ}$. *By definitions 1,2,* $(x)^{\bar{F}} \subseteq (x)$, *then* $\alpha \subseteq \alpha^F$. *Apparently, there exists difference attribute set between* α^F *and* α, $\nabla\alpha^F = \{\beta_i \mid \alpha_i \in \alpha^F, \bar{f}(\alpha_i) = \beta_i \bar{\in} \alpha^F, \bar{f} \in \bar{F}\}$. *If* $(x)^{\bar{F}}$ *is restored to* (x), *or* $(x)^{\bar{F}} = (x)$, *then* $(x)^{\bar{F}}$ *and* (x) *have common attribute set, or* $\nabla\alpha^F$ *is deleted from* α^F , *then there is formula (24).* $\mathbf{2}^{\circ}$. *If* $\nabla\alpha^F = \{\beta_i \mid \alpha_i \in \alpha^F, \bar{f}(\alpha_i) = \beta_i \bar{\in} \alpha^F, \bar{f} \in \bar{F}\}$ is deleted from α^F , or $\alpha^F - \{\beta_i \mid \alpha_i \in \alpha^F, \bar{f}(\alpha_i) = \beta_i \bar{\in} \alpha^F, \bar{f} \in \bar{F}\} = \alpha$,then $(x)^{\bar{F}}$ and (x) have common attribute set , then $(x)^{\bar{F}} = (x)$. So $(x)^{\bar{F}}$ is restored to (x).

Theorem 5. *(Recovery theorem of attribute of* $\bar{F}$ *-compressed data sequence)*
In the sequence of $\bar{F}$ *-compressed data* $(x)_i^{\bar{F}}$, *the necessary and sufficient condition of* $(x)_i^{\bar{F}}$ *being restored to data* $(x)_{i-1}^{\bar{F}}$ *is: the attribute set* α_i^F *of* $(x)_i^{\bar{F}}$ *and attribute set* α_{i-1}^F *of* $(x)_{i-1}^{\bar{F}}$ *satisfy*

$$\alpha_i^F - \{\beta_j \mid \alpha_j \in \alpha_i^F, \bar{f}(\alpha_j) = \beta_j \,\bar{\in}\, \alpha_i^F, \bar{f} \in \bar{F}\} = \alpha_{i-1}^F. \tag{25}$$

where $i = n, n-1, \cdots, 1$.

The proof is similar to the proof of theorem 4. It is omitted.

By theorem 4,5, we get:

Proposition 2. *Deleting redundant attributes of $\bar{F}$-compressed data, the data is restored; and vice versa.*

Theorem 6. *(Recovery theorem of data element of $\bar{F}$-compressed data) The necessary and sufficient condition of $\bar{F}$-compressed data $(x)^{\bar{F}}$ being restored to data (x) is that there exists transfer function $f \in F$ and $(x)^{\bar{F}}$, (x) satisfy*

$$(x)^{\bar{F}} \cup \{x_i \mid x_i \in (x)^-, x_i \,\bar{\in}\, (x)^{\bar{F}}, f(x_i) = x_i' \in (x)^{\bar{F}}, f \in F\} = (x). \tag{26}$$

Proof. Suppose that data $(x) = \{x_1, x_2, \cdots, x_q\}$ is compressed to $\bar{F}$-compressed data $(x)^{\bar{F}} = \{x_1, x_2, \cdots, x_p\}$, then there exists transfer function $\bar{f} \in \bar{F}$, making $\bar{F}$-redundant data $(x)^- = \{x_{p+1}, x_{p+2}, \cdots, x_q\} = \{x_i \mid x_i \in (x), \bar{f}(x_i) = x_i' \,\bar{\in}\, (x), \bar{f} \in \bar{F}\}$. **1°.** If $(x)^{\bar{F}}$ is restored to data (x), then $(x)^{\bar{F}} = (x)$. By theorem 4, the attribute set α^F of $(x)^{\bar{F}}$ and attribute set α of (x) satisfy formula (24), or deleting attributes from α^F to get formula (24), then adding data element x_i to $(x)^{\bar{F}}$ to get $(x)^{\bar{F}} = (x)$. By definitions 1,2, $x_i \in (x)^-$. So, there exists transfer function $f \in F$, making formula (26). **2°.** If there exists transfer function $f \in F$, it satisfies formula (26), or, adding $\{x_i \mid x_i \in (x)^-, x_i \,\bar{\in}\, (x)^{\bar{F}}, f(x_i) = x_i' \in (x)^{\bar{F}}, f \in F\}$ to $(x)^{\bar{F}}$, then $(x)^{\bar{F}} = (x)$, or $(x)^{\bar{F}}$ is restored to (x).

Theorem 7. *(Recovery theorem of data element $\bar{F}$-compressed data sequence) The necessary and sufficient condition of $(x)_i^{\bar{F}}$ being restored to $(x)_{i-1}^{\bar{F}}$ in the sequence of $\bar{F}$-compressed data sequence $(x)_i^{\bar{F}}$ is: there exists transfer function $f_i \in F$, $(x)_i^{\bar{F}}$ and $(x)_{i-1}^{\bar{F}}$ satisfy*

$$(x)_i^{\bar{F}} \cup \{x_k \mid x_k \in (x)_i^-, x_k \,\bar{\in}\, (x)_i^{\bar{F}}, f(x_k) = x_k' \in (x)_i^{\bar{F}}, f_i \in F\} = (x)_{i-1}^{\bar{F}}, \tag{27}$$

where $i = n, n-1, \cdots, 1$.

The proof is similar to the proof of theorem 6. It is omitted.

Theorems 5, 7, can be expanded to recovery between non-adjacent $\bar{F}$-compressed data $(x)_j^{\bar{F}}$ and $(x)_i^{\bar{F}}$.

From definition 4, theorems 3, 5, 7, there is a fact as this: $\bar{F}$-compression and $\bar{F}$-recovery are two opposite processes. Starting from the data (x), $\bar{F}$-compression proceeds step by step, until the compressed data is obtained and satisfies formula (19). Starting from the highest compressed data, $\bar{F}$-recovery proceeds step by step, until the data is obtained. But it can to be restored in leap-frog. The discussion about data recovery degree is in section 4.

4 Measures of $\bar{F}$-Compressed Data and $\bar{F}$-Redundant Data

Definition 1. *$\xi_j^{\bar{F}}$ is called $\bar{F}$-compressed degree of $\bar{F}$-compressed data $(x)_j^{\bar{F}}$ with respect to data $(x)_i^{\bar{F}}$, called $\bar{F}$-compressed degree of data $(x)_j^{\bar{F}}$ for short, if*

$$\xi_j^{\bar{F}} = \mathbf{card}((x)_j^{\bar{F}}) / \mathbf{card}((x)_i^{\bar{F}}), \tag{28}$$

where card=cardinal number.

Definition 2. *$\psi_j^{\bar{F}}$ is called $\bar{F}$-redundant degree of $\bar{F}$-redundant data $(x)_j^-$ with respect to $\bar{F}$-compressed data $(x)_i^{\bar{F}}$, called $\bar{F}$-redundant degree of $(x)_j^-$ for short, if*

$$\psi_j^{\bar{F}} = \mathbf{card}((x)_j^-) / \mathbf{card}((x)_i^{\bar{F}}). \tag{29}$$

Definition 3. *$\zeta_j^{\bar{F}}$ is called $\bar{F}$-recovery degree of $\bar{F}$-compressed data $(x)_j^{\bar{F}}$ with respect to data $(x)_i^{\bar{F}}$, called $\bar{F}$-recovery degree of data $(x)_j^{\bar{F}}$ for short, if*

$$\zeta_j^{\bar{F}} = \mathbf{card}((\alpha)_i^F) / \mathbf{card}((\alpha)_j^F), \tag{30}$$

where $(\alpha)_i^F$, $(\alpha)_j^F$ are attribute sets of data $(x)_i^{\bar{F}}$ and data $(x)_j^{\bar{F}}$ respectively.

By formulas (28) and (29), we can easily get:

Theorem 1. *(Relation theorem of $\bar{F}$-compressed degree and $\bar{F}$-redundant degree) If $\xi_i^{\bar{F}}$ and $\psi_i^{\bar{F}}$ are $\bar{F}$-compressed degree and $\bar{F}$-redundant degree of $\bar{F}$-compressed data $(x)_{i-1}^{\bar{F}}$ respectively, then*

$$\xi_i^{\bar{F}} + \psi_i^{\bar{F}} = 1, \tag{31}$$

where $i = n, n-1, \cdots, 1$.

In fact, By definitions 1,2, $\xi_i^{\bar{F}} + \psi_i^{\bar{F}} = \mathbf{card}((x)_i^{\bar{F}}) / \mathbf{card}((x)_{i-1}^{\bar{F}}) + \mathbf{card}((x)_i^-) / \mathbf{card}((x)_{i-1}^{\bar{F}})$ $= (\mathbf{card}((x)_i^{\bar{F}}) + \mathbf{card}((x)_i^-)) / \mathbf{card}((x)_{i-1}^{\bar{F}}) = \mathbf{card}((x)_{i-1}^{\bar{F}}) / \mathbf{card}((x)_{i-1}^{\bar{F}}) = 1$.

By definition 3, we get :

Theorem 2. *(Compressed data recovery theorem of $\bar{F}$-recovery degree) In the sequence of $\bar{F}$-compressed data for data (x), the necessary and sufficient condition of $(x)_j^{\bar{F}}$ restored to data $(x)_i^{\bar{F}}$ is: $\bar{F}$-recovery degree of $(x)_j^{\bar{F}}$ with respect to $(x)_i^{\bar{F}}$ satisfies $\zeta_j^{\bar{F}} = 1$.*

In fact, $(x)_j^{\bar{F}}$ can be restored to $(x)_i^{\bar{F}}$, if and only if there is: deleting redundant attributes from attribute set α_j^F of $(x)_j^{\bar{F}}$ to make α_j^F and attribute set α_i^F of $(x)_i^{\bar{F}}$ satisfy: $\alpha_j^F - \{\beta_k \mid \alpha_k \in \alpha_j^F, \bar{f}(\alpha_k) = \beta_k \bar{\in} \alpha_j^F, \bar{f} \in \bar{F}\} = \alpha_i^F$, or $\alpha_j^F = \alpha_i^F$. From formula (30), there is $\zeta_j^{\bar{F}} = \mathbf{card}((\alpha)_i^F) / \mathbf{card}((\alpha)_j^F) = 1$.

Corollary 1. *The $\bar{F}$-recovery degree of data $(x)_j^{\bar{F}}$ with respect to data $(x)_i^{\bar{F}}$ the bigger, the recovery effect of $(x)_j^{\bar{F}}$ the better, and vice versa.*

5 Application of $\bar{F}$-Compressed Data and $\bar{F}$-Redundant Data in Communication System

Assumption 1. $y=\{y_1,y_2,\cdots,y_q\}$ *is characteristic value set of data* $(x)=\{x_1,x_2,\cdots,x_q\}$. *For simplicity, not misleading, in this section,* y *is called data;* $y_1^{\bar{F}}$, y_1^{-} *are* $\bar{F}$ *-compressed data and* $\bar{F}$ *-redundant data of* y *respectively;* $y_2^{\bar{F}}$, y_2^{-} *are* $\bar{F}$ *-compressed data and* $\bar{F}$ *-redundant data of* $y_1^{\bar{F}}$ *respectively; specially,* $y_\circ^{\bar{F}}=y$, $y_\circ^{-}=\phi$ *are initial values of compressed data and redundant data of* y.

The example of this section comes from experimental data sample of Natural Science Foundation of Fujian (N_o.2009J01294).

Table 1. Redundant data y_o^{-}, y_1^{-}, y_2^{-} in the transmission of data

y	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8
y_o^{-}	–	–	–	–	–	–	–	–
y_1^{-}	–	1.53	1.12	–	1.28	–	–	–
y_2^{-}	–	–	–	1.37	–	1.91	1.34	–

Table 2. Compressed data $y_o^{\bar{F}}$, $y_1^{\bar{F}}$, $y_2^{\bar{F}}$ in the transmission of data

y	y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8
$y_o^{\bar{F}}=y$	1.62	1.53	1.12	1.37	1.28	1.91	1.34	1.47
$y_1^{\bar{F}}$	1.62	–	–	1.37	–	1.91	1.34	1.47
$y_2^{\bar{F}}$	1.62	–	–	–	–	–	–	1.47

In table 1, table 2, the data is from raw data which is processed through the technical means on the typhoon pre-warning and search-rescue system. It does not affect the analysis of results. In table 1, 2, "–" means "null data", data $y=\{y_1,y_2,y_3,y_4,y_5,y_6,y_7,y_8\}$ is initial data of the system transmission, $y_\circ^{\bar{F}}=y$ and $y_\circ^{-}=\phi$ are initial states of $\bar{F}$-compressed data and $\bar{F}$-redundant data before data y is compressed. Adding attributes to attribute set α of y to get α_1^{F}, deleting redundant data $y_1^{-}=\{y_2,y_3,y_5\}$ from y, thus compressed data $y_1^{\bar{F}}=\{y_1,y_4,y_6,y_7,y_8\}$ is obtained. Adding attributes to attribute set α_1^{F} of $y_1^{\bar{F}}$ to get α_2^{F}, deleting redundant data $y_2^{-}=\{y_4,y_6,y_7\}$ from $y_1^{\bar{F}}$, the compressed data $y_2^{\bar{F}}=\{y_1,y_8\}$ of $y_1^{\bar{F}}$ is obtained, moreover $y_2^{\bar{F}}\subseteq y_1^{\bar{F}}\subseteq y$. By formula (22), $y_2^{\bar{F}}$, $y_1^{\bar{F}}$, y can be identified ; By formula (20), y one-directionally depends on $y_1^{\bar{F}}$, $y_1^{\bar{F}}$ one-directionally depends on $y_2^{\bar{F}}$, or $y_2^{\bar{F}}\Rightarrow y_1^{\bar{F}}\Rightarrow y$. Redundant data y_1^{-}, y_2^{-} and compressed data $y_1^{\bar{F}}$, $y_2^{\bar{F}}$ satisfy formula (15) in theorem 1 for section 3.

Using theorems 5,4, if the redundant attributes in α_2^F , α_1^F are deleted, then there is (25); or if the redundant data is supplemented , then there is formula (27). So data $y_2^{\bar{F}}$ is restored to data $y_1^{\bar{F}}$, data $y_1^{\bar{F}}$ is restored to data y . The compressed degrees of $y_1^{\bar{F}}$ and $y_2^{\bar{F}}$ are 0.625 and 0.4 respectively; the redundant degrees of $y_1^{\bar{F}}$ and $y_2^{\bar{F}}$ are 0.375 and 0.6 respectively. They satisfy formula (31) in theorem 1 for section 4. In our research project "The study of typhoon pre-warning and search-rescue system based on GPS/3G/GIS on the sea in Ningde ", the interrelated applications of $\bar{F}$ -compressed data and $\bar{F}$ -redundant data have been confirmed in both data compression collected by sensor and data recovery by base station.

6 Discussion

The paper uses deleted sets and internal P-sets of P-sets to study redundant data and compressed data in data processing. The concepts of $\bar{F}$ -compressed data and $\bar{F}$ -redundant data are given. Compressed-redundant data pair $((x)^{\bar{F}},(x)^{-})$ is obtained.

Using dynamic characteristic of P-sets, the dynamic compressed data sequence and dynamic redundant data sequence are obtained. In the process of generation and recovery of $\bar{F}$ -compressed data, the generation and recovery of $\bar{F}$ -redundant data have been resolved because $\bar{F}$ -compressed data and $\bar{F}$ -redundant data are complementary each other. In addition, the compressed data sequence one-directionally depends on each other sequentially and can be identified. In the existing literatures, this paper applies firstly the deleted sets of P-sets to discuss the redundant data in compressing data, so that the applications of P-sets are expanded. In fact, the supplemented sets of P-sets can also be applied to discuss the redundant data in data expansion to get more results. From the discussion, we can see: P-sets are a mathematical tool and method to study dynamic information system; P-sets have great potential and space of application in the fields which have dynamic characteristics.

Acknowledgements. This work was supported by the Natural Foundation of Fujian Province of China (2009J01294) and the Natural Science Foundation of Shandong Province of China (Y2007H02), as well as by the Ningde Normal University Foundation of China (2008J002).

References

1. Shi, K.Q.: P-sets and its applications. An International Journal Advances in Systems Science and Application 9(2), 209–219 (2009)
2. Shi, K.Q.: P-sets. Journal of Shandong University (Natural Science) 43(11), 77–84 (2008) (in Chinese)
3. Shi, K.Q., Zhang, L.: Internal P-sets and data outer-recovery. Journal of Shandong University (Natural Science) 44(4), 8–14 (2009) (in Chinese)

4. Li, Y.Y.: Generation and Recovery of *F*-distortion Data. Journal of Jishou University (Natural Science) 31(3), 59-72 (2010) (in Chinese)
5. Yu, X.Q.: Identification and filtering of P-sets. Journal of Shandong University (Natural Science) 45(1), 94–98 (2010) (in Chinese)
6. Yu, X.Q.: Area characteristic of P-sets. Journal of Shandong University (Natural Science) 45(2), 99–102 (2010) (in Chinese)
7. Zhang, F., Chen, P., Zhang, L.: P-separation of P-sets and its applications. Journal of Shandong University (Natural Science) 45(3), 71–75 (2010) (in Chinese)
8. Shi K.Q., Li X.H.: Camouflaged information identification and its applications. An International Journal Advances in Systems Science and Applications, 10(2), 157–167 (2010)
9. Lin H.K., Li Y.Y.: P-sets and its P-separation theorems. An International Journal Advances in Systems Science and Applications, 10(2), 209-215(2010)
10. Shi, K.Q., Yao, B.X.: Function S-rough sets and systems law identification. Sciences in China (F) 51(5), 499–510 (2008)
11. Shi, K.Q., Zhao, J.L.: Function S-rough sets and Security-authentification of hiding law. Sciences in China (F) 51(7), 924–935 (2008)
12. Shi, K.Q., Yao, B.X.: Function S-rough sets and systems law mining, pp. 51–68. Science Press, China (2007)

On the Decomposition of Distribution Function by Using Hausdorff Calculus

Ying-jian Lin and Shi-pan Lu

Department of Mathematics, School of Science,
Jimei University, Xiamen, Fujian 361021, P.R. China
e-mail: linyj@jmu.edu.cn

Abstract. This paper, by using Hausdorff calculus theory, decomposes farther a singular continuous distribution function into a sum of a series of absolute continuous distribution functions in different levels with respect to the Hausdorff measures and a singular continuous distribution function with respect to the Hausdorff measures. Consequently, it gives a more accurate decomposition formula for the distribution functions.

Keywords: Hausdorff measure, Hausdorff fractional derivation, $\mathcal{H}^s$ - absolute continuous, $\mathcal{H}^s$ - singular continuous, distribution function.

1 Introduction

A real function F is called a *distribution function*, if F is a right continuous monotone increasing function, and satisfies $F(-\infty)=0$ and $F(+\infty)=1$. Chung K. L. gives in [1].

Theorem 1. *Every distribution function F can be expressed as a convex combination of a discrete function F_d, an absolute continuous function F_{ac}, and a singular continuous distribution function F_s, i.e.*

$$F=\alpha_1 F_d+\alpha_2 F_{ac}+\alpha_3 F_s ,$$

where $\alpha_1+\alpha_2+\alpha_3=1$, $\alpha_1>0$, $\alpha_2>0$, $\alpha_3>0$, and this decomposition is uniform.

In this paper, by using Hausdorff calculus theory, we decompose farther the singular continuous distribution function into a sum of a series of absolute continuous distribution functions in different levels with respect to the Hausdorff measures and a singular continuous distribution function with respect to the Hausdorff measure, and then gives a more accurate decomposition formula for the distribution functions.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 673–681.
springerlink.com

2 Basic Concepts

We assume that the readers are familiar with the definition and the properties of the Hausdroff measure, otherwise the details can be found in [2] and [3]. In this paper, we always assume the set $E \subset (-\infty,+\infty)$.

A sequence of open intervals $\{I_i\}$ is said to be a *cover* of E , if $E \subset \bigcup_i I_i$; let $\delta > 0$, $\{I_i\}$ is said to be a $\delta-$ *cover* of E if $0 < |I_i| \le \delta$ for each i .

Let $0 < s \le 1$. The $s-$ *dimensional Hausdorff measure* is defined as

$$\mathcal{H}^s(E) = \lim_{\delta \to 0} \inf_{0<|I_i|\le\delta} \Sigma_i |I_i|^s ,$$

where the infimum is taken over all $\delta-$ cover of E .

It is easily seen that $\mathcal{H}^1 = \mathcal{L}$, the *Lebesgue measure*, when $s = 1$.

A set E is called a $s-$ *set*, if E is $\mathcal{H}^s-$ measurable and $o < \mathcal{H}^s(E) < +\infty$.

A collection of sets $\mathcal{V}$ is called a *Vitali class* for E , if for each $x \in E$ and $\delta > 0$, there exists $V \in \mathcal{V}$ with $x \in V$ and $0 < |V| \le \delta$.

Theorem 2. *Let $E \subset (-\infty,+\infty)$ be $\mathcal{H}^s-$ measurable, $0 < s \le 1$, and let $\mathcal{V} = \{I\}$ be a Vitali class of closed intervals of E , then we may select a countable or finite sequence of non-overlapping intervals $\{I_i\}$ from $\mathcal{V}$, such that either $\mathcal{H}^s(E \setminus \bigcup_i I_i) = 0$ or $\Sigma_i |I_i| = \infty$. Further, if $\mathcal{H}^s(E) < \infty$, then, for given $\varepsilon > 0$, we may also require that $\mathcal{H}^s(E) \le \Sigma_i |I_i|^s + \varepsilon$.*

Let $\mathcal{F}$ be a collection of increasing right continuous functions on $(-\infty,+\infty)$, $F \in \mathcal{F}$, and $0 < s \le 1$.

The *upper s-derivate* of function F at $x \in (-\infty,+\infty)$ is defined by

$$\bar{D}^s F(x) = \lim_{\delta \to 0} \sup_{x \in I, |I| < \delta} \frac{F(I)}{|I|^s} ,$$

where I is an interval. When $s = 1$, it is the ordinary upper derivate of F at x , and denote it by $\bar{D}F(x)$.

A function $F \in \mathcal{F}$ is said to be *absolute continuous with respect to* $\mathcal{H}^s$, or $\mathcal{H}^s - AC$ for short, if for any $\varepsilon > 0$ there is a number $\eta > 0$ so that for any sequence of closed intervals $\{I_i\}$, $\Sigma_i F(I_i) < \varepsilon$ whenever $\Sigma_i |I_i|^s < \eta$. When $s = 1$, F is *Lebesgue absolute continuous*, or $\mathcal{L} - AC$ for short.

Obviously, if F is $\mathcal{H}^t - AC$ and $0 < s < t \le 1$, then F is $\mathcal{H}^s - AC$, and the contrary is not.

A function $F \in \mathcal{F}$ is said to be *singular with respect to* $\mathcal{H}^s$, or $\mathcal{H}^s -$ *singular* for short, if $F \not\equiv 0$, and there is a set $E \subset (-\infty,+\infty)$ with $\mathcal{H}^s(E) = 0$ so that for each

$x \in E^c$ we have $\bar{D}^s F(x) = 0$. Especially, when $s = 1$, F is *Lebesgue singular,* or $\mathcal{L}-$ *singular* for short.

Let $E \subset (-\infty, +\infty)$ be an $s-$set, f a function of $\mathcal{H}^s-$measurable non-negative on E. Then the *Hausdorff integral* of f over E is defined to be

$$\int_E f \mathrm{d}\mathcal{H}^s = \sup \sum_i v_i \mathcal{H}^s(E_i),$$

where the supremum is taken over all finite sequence $\{E_i\}$ of disjoint $\mathcal{H}^s-$ measurable sets with $E = \bigcup_i E_i$, and for each i, $v_i = \inf\{f(x) : x \in E_i\}$. Especially, when $s = 1$, it is *Lebesgue integral*, write it as $\int_E f \mathrm{d}\mathcal{L}$ (see **[7]**).

Hausdorff integral has the basic properties as usual Lebesgue integral dose, for example, $\int_E f \mathrm{d}\mathcal{H}^s$ is $\mathcal{H}^s - AC$ on E.

Let $F \in \mathcal{F}$, the increment of F over an interval $I = (u, v)$ or $[u, v] \subset (-\infty, +\infty)$ is the difference $F(v) - F(u)$. We often write it as $F(I)$ or $F(v, u)$. For $E \subset (-\infty, +\infty)$, define the total variational of F on E as

$$F^*(E) = \inf \sum_i F(I_i),$$

where the infimum is taken over all covers $\{I_i\}$ of E. We can easily verify that $F^*(I) = F(I)$ for any interval I.

3 Main Results

In this section, we will restrict $F \in \mathcal{F}$, $E \subset (-\infty, +\infty)$, and $0 < s \le 1$.

Lemma 1. [6] *Let* $E \subset (-\infty, +\infty)$ *be an* $s-$*set.*

(i) If there is $c > 0$ *so that* $\bar{D}^s F(x) \le c$ *for each* $x \in E$*, then* $F^*(E) \le c\mathcal{H}^s(E)$*;*

(ii) If there is $c > 0$ *so that* $\bar{D}^s F(x) \ge c$ *for each* $x \in E$*, then* $F^*(E) \ge c\mathcal{H}^s(E)$*;*

(iii) If $\bar{D}^s F(x) = 0$ *for all* $x \in E$*, then* $F^*(E) = 0$*;*

(iv) If $E_\infty = \{x \in E : \bar{D}^s F(x) = +\infty\}$*, then* $\mathcal{H}^s(E_\infty) = 0$. □

Lemma 2. [6]. *Let* $E \subset (-\infty, +\infty)$ *be an* $s-$*set, then we have*

$$F^*(E \setminus E_\infty) = \int_E \bar{D}_s F \mathrm{d}\mathcal{H}^s,$$

where $E_\infty = \{x \in E : \bar{D}^s F(x) = +\infty\}$. □

Theorem 3. *Suppose that* F *is an* $\mathcal{H}^{s_0}-$ *singular continuous distribution function* ($0<s_0\le 1$). *Let* $1\ge s_0>s_1>\cdots>s_n>0$, $E_i=\{x\in E_{i-1}:\bar{D}^{s_{i-1}}F(x)=\infty\}$, $i=1,2,\cdots,n+1$, *and* E_i *an* s_i- *set,* $i=1,2,\cdots,n$, *where* $E_0=(-\infty,\infty)$, *then we have the decomposition*

$$F=\sum_{i=1}^{n+1}\alpha_i F_i\ ,\ \sum_{i=1}^{n+1}\alpha_i=1\ ,\ \alpha_i>0\ ,\ i=1,2,\cdots,n+1\ ,$$

where F_i *is an* $\mathcal{H}^{s_i}-AC$ *distribution function with* $F_i(x)=\alpha_i^{-1}\int_{(-\infty,x]\cap E_i}\bar{D}^{s_i}F\mathrm{d}\mathcal{H}^{s_i}$, $i=1,2,\ \cdots,n$, F_{n+1} *is an* $\mathcal{H}^{s_n}-$ *singular continuous distribution function.*

Proof. Let

$$E_i^*=\{x\in E_i:\bar{D}^{s_i}F(x)<\infty\}\ ,\ i=0,1,\cdots,n\ .$$

Then $E_{i-1}=E_{i-1}^*\cup E_i$, $i=1,2,\cdots,n+1$, $E_i^*\cap E_j^*=\phi$ ($i\ne j$), $E_0=\bigcup_{i=0}^{n}E_i^*\cup E_{n+1}$. Write

$$G_i(x)=F^*((-\infty,x]\cap E_i^*)\ ,\ i=0,1,\cdots,n\ ,$$

and

$$G_{n+1}(x)=F^*((-\infty,x]\cap E_{n+1})\ ,$$

then

$$F=\sum_{i=0}^{n+1}G_i\ .$$

By Lemma 2, we have

$$G_i(x)=F^*((-\infty,x]\cap E_i^*)=\int_{(-\infty,x]\cap E_i}\bar{D}^{s_i}F\mathrm{d}\mathcal{H}^{s_i}\ ,$$

and G_i is $\mathcal{H}^{s_i}-AC$, $i=0,1,\cdots,n$. By Lemma 1 (iv), we see that $\mathcal{H}^{s_n}(E_{n+1})=0$, and then G_{n+1} is of $\mathcal{H}^{s_n}-$ singular. Furthermore, since F and G_i ($i=0,1,\cdots,n$) are all continuous, so G_{n+1} is also continuous.

Notice that F is $\mathcal{H}^{s_0}-$ singular, we see that there is $D\subset(-\infty,\infty)$, which satisfying $\mathcal{H}^{s_0}(D)=0$, such that $\bar{D}^{s_0}F(x)=0$ for all $x\in D^c$, therefore we have

$$G_0(x)=\int_{(-\infty,x]}\bar{D}^{s_0}F\mathrm{d}\mathcal{H}^{s_0}\equiv 0\ .$$

Evidently, since $\bar{D}^{s_i}F(x)\ge 0$ for any $x\in E_i$, $i=1,2,\cdots,n$, we see that each G_i ($i=1,2,\ \cdots,n$) is monotone increasing. Therefore, it follows from

$$\sum_{i=1}^{n}G_i(x)-\sum_{i=1}^{n}G_i(x')=\sum_{i=1}^{n}F^*([x',x]\cap E_i^*)$$

$$= F^*([x',x] \cap \bigcup_{i=1}^{n} E_i^*) \le F^*([x',x]) = F(x) - F(x')$$

that

$$G_{n+1}(x) - G_{n+1}(x') = [F(x) - F(x')] - [\sum_{i=1}^{n} G_i(x) - \sum_{i=1}^{n} G_i(x')] \ge 0 ,$$

i.e. G_{n+1} is also monotone increasing.

Forthermore, since $F(-\infty) = 0$ and $F(\infty) = 1$, we have $0 \le G_i \le 1$ for each $i = 1, 2, \cdots, n+1$. Let $0 < G_i(\infty) < 1$ for each $i = 1, 2, \cdots, n+1$, write $\alpha_i = G_i(\infty)$ and let

$$F_i(x) = \alpha_i^{-1} G_i(x) , \ x \in (-\infty, \infty) , \ i = 1, 2, \cdots, n+1 ,$$

then for each $i = 1, 2, \cdots, n$, F_i is an $\mathcal{H}^{s_i} - AC$ distribution function with

$$F_i(x) = \alpha_i^{-1} \int_{(-\infty,x] \cap E_i} \overline{D}^{s_i} F \mathrm{d}\mathcal{H}^{s_i} ,$$

at the same time F_{n+1} is an $\mathcal{H}^{s_n}$ − singular continuous distribution function, and we have

$$F = \sum_{i=1}^{n+1} \alpha_i F_i ,$$

where $\Sigma_{i=1}^{n+1} \alpha_i = \Sigma_{i=1}^{n+1} G_i(\infty) = F(\infty) = 1$, $\alpha_i > 0$, $i = 1, 2, \cdots, n+1$, and the theorem follows. □

Since the Theorem 3, the Theorem 1 can be improved as following:

Theorem 4. *Every distribution function F can be expressed as a convex combination of a discrete distribution function, an absolute continuous distribution function, a series of absolute continuous distribution functions in different levels with respect to the Hausdorff measures, and a singular continuous distribution function with respect to the Hausdorff measures. That is, if $1 = s_0 > s_1 > \cdots > s_n > 0$, $E_0 = (-\infty, \infty)$, such that $E_i = \{x \in E_{i-1} : \overline{D}^{s_{i-1}} F(x) = \infty\}$, $i = 1, 2, \cdots, n+1$, and E_i is an $s_i -$ set, $i = 1, 2, \cdots, n$, then we have the decomposition formula*

$$F = \alpha_d F_d + \sum_{i=0}^{n+1} \alpha_i F_i , \ \alpha_d + \sum_{i=0}^{n+1} \alpha_i = 1 , \ \alpha_d > 0 , \ \alpha_i > 0 , \ i = 0, 1, 2, \cdots, n+1 ,$$

where F_d is a discrete distribution function, F_0 is an $\mathcal{L} - AC$ distribution function, F_i with $F_i(x) = \alpha_i^{-1} \int_{(-\infty,x] \cap E_i} \overline{D}^{s_i} F \mathrm{d}\mathcal{H}^{s_i}$ is an $\mathcal{H}^{s_i} - AC$ distribution function, $i = 1, 2, \cdots, n$, and F_{n+1} is an $\mathcal{H}^{s_n}$ − singular continuous distribution function.

4 An Example

Here we give a decomposition example of the singular continuous distribution function.

Firstly, we construct a Cantor sets E_1 and a Cantor-like set E_2 :

By separating the unite closed interval $[0,1]$ into three parts with equal length $1/3$ marked as $\Delta_i=[i/3,(i+1)/3]$, $i=0,1,2$, and removing the middle open interval Δ_1^o, it leaves two closed intervals Δ_0 and Δ_2. Similarly, for each $\lambda_1\in\{0,2\}$, by removing the middle open interval $\Delta_{\lambda_1 1}^o$ from Δ_{λ_1}, leaves two closed intervals $\Delta_{\lambda_1 0}$ and $\Delta_{\lambda_1 2}$. Carrying on this procedure, we obtain a sequence of open intervals $\Delta^o{}_{\lambda_1\cdots\lambda_{n-1}1}$ ($\lambda_i=0,2$) and a sequence of closed intervals $\Delta_{\lambda_1\cdots\lambda_n}$ ($\lambda_i=0,2$). Write

$$E_1=\bigcap_n \bigcup_{\substack{\lambda_i=0,\text{or}2\\ i=1,2,\cdots,n}} \Delta_{\lambda_1\cdots\lambda_n} ,$$

it is easily to see that each $x\in E_1$ can be expressed as

$$x=\sum_{i=1}^{\infty}\lambda_i 3^{-i} ,\ \ \lambda_i=0, \text{ or } 2 .$$

The set E_2 is constructed as follows: by separating the unite closed interval $[0,1]$ into five parts with equal length $1/5$ marked as $\Delta_i=[i/5,(i+1)/5]$, $i=0,1,2,3,4$, and removing two open intervals $\Delta_{\theta_1}^o$, $\theta_1=1,3$, it leaves three closed intervals Δ_{ε_1}, $\varepsilon_1=0,2,4$. Similarly, for each $\varepsilon_1\in\{0,2,4\}$, by removing two open intervals $\Delta_{\varepsilon_1\theta_2}^o$ from Δ_{ε_1}, $\theta_2=1,3$, leaves three closed intervals $\Delta_{\varepsilon_1\varepsilon_2}$, $\varepsilon_2=0,2,4$. Carrying on this procedure, we obtain a sequence of open intervals $\Delta^o{}_{\varepsilon_1\cdots\varepsilon_{n-1}\theta_n}$ ($\varepsilon_i=0,2,4$; $\theta_n=1,3$) and a sequence of closed intervals $\Delta_{\varepsilon_1\cdots\varepsilon_n}$ ($\varepsilon_i=0,2,4$). Write

$$E_2=\bigcap_n \bigcup_{\substack{\varepsilon_i=0,2,4\\ i=1,2,\cdots,n}} \Delta_{\varepsilon_1\cdots\varepsilon_n} ,$$

we see that each $x\in E_2$ can be expressed as

$$x=\sum_{i=1}^{\infty}\varepsilon_i 5^{-i} ,\ \ \varepsilon_i=0,2, \text{ or } 4 .$$

It is easily to check that $mE_1=0$ and $mE_2=0$, but $\mathcal{H}^{s_1}(E_1)=1$ and $\mathcal{H}^{s_2}(E_2)=1$, where

$$s_1=\dim_{\mathcal{H}} E_1=\log 2/\log 3\doteq 0.630929753 ,$$
$$s_2=\dim_{\mathcal{H}} E_2=\log 3/\log 5\doteq 0.682606194 .$$

Secondly, we define two functions $F_1,\ F_2:[0,1]\to(-\infty,\infty)$ as follows

$$F_1(x)=\sum_{i=1}^{n-1}\frac{\lambda_i}{2^{i-1}}+\frac{1}{2^n} ,\qquad x\in\Delta^o{}_{\lambda_1\cdots\lambda_{n-1}1} ,\ \lambda_i=0,2 ,$$
$$=\sum_{i=1}^{\infty}\frac{\lambda_i}{2^{i-1}} ,\qquad x=\sum_{i=1}^{\infty}\frac{\lambda_i}{3^{i-1}}\in E_1 ,\ \lambda_i=0,2 ,$$

$$
\begin{aligned}
&= 0\,, && x \le 0\,, \\
&= 1\,, && x \ge 1\,; \\
F_2(x) &= \frac{1}{2}\sum_{i=1}^{n-1}\frac{\varepsilon_i}{3^i} + \frac{1}{3^n}\,, && x \in \Delta^o_{\varepsilon_1\cdots\varepsilon_{n-1}1}\,,\ \varepsilon_i = 0,2,4\,, \\
&= \frac{1}{2}\sum_{i=1}^{n-1}\frac{\varepsilon_i}{3^i} + \frac{2}{3^n}\,, && x \in \Delta^o_{\varepsilon_1\cdots\varepsilon_{n-1}3}\,,\ \varepsilon_i = 0,2,4\,, \\
&= \frac{1}{2}\sum_{i=1}^{\infty}\frac{\varepsilon_i}{3^i}\,, && x = \sum_{i=1}^{\infty}\frac{\varepsilon_i}{5^i} \in E_2\,,\ \varepsilon_i = 0,2,4\,, \\
&= 0\,, && x \le 0\,, \\
&= 1\,, && x \ge 1\,.
\end{aligned}
$$

Let

$$F(x) = \frac{1}{2}F_1(x) + \frac{1}{2}F_2(x)\,.$$

We see that $F(-\infty)=0$, $F(\infty)=1$. It is easily to check that F_1 and F_2 are all monotone increasing continuous in $(-\infty,\infty)$, and

$$\bar{D}F_i(x) = 0\,,\ x \in [0,1]\setminus E_i\,,\ i = 1,2\,,$$

thus F is also monotone increasing continuous in $(-\infty,\infty)$, and $\bar{D}F(x)=0$, $x\in[0,1]$ $\setminus(E_1\cup E_2)$. Therefore F is a singular continuous distribution function in $(-\infty,\infty)$. Further -more, it can be proved that

$$\bar{D}^{s_i}F_i(x) = 1\,,\ x \in E_i\,,\ i = 1,2\,,$$

therefore F_i is $\mathcal{H}^{s_i}-AC$ in $(-\infty,\infty)$ respectively, $i=1,2$. It follows that the singular continuous distribution function F can be expressed as a convex combination of the $\mathcal{H}^{s_1}-AC$ distribution function F_1 and the $\mathcal{H}^{s_2}-AC$ distribution function F_2.

Thirdly, we will only show that $\bar{D}^{s_2}F_2(x)=1$ for $x\in E_2$, where $s_2=\log 3/\log 5$, and for convenience, will omit the index 2.

Let $x\in E$. Then x can be expressed as $x=\Sigma_{i=1}^{\infty}\varepsilon_i 5^{-i}$, where $\varepsilon_i=0,2,4$, $i=1,2,\cdots$, and then $x\in\Delta_{\varepsilon_1\cdots\varepsilon_n}$ for each positive integer n. By the fact

$$\frac{F(\Delta_{\varepsilon_1\cdots\varepsilon_n})}{\left|\Delta_{\varepsilon_1\cdots\varepsilon_n}\right|^s} = \frac{3^{-n}}{(5^{-n})^s} = 1\,,$$

we see that $\bar{D}^sF(x)\ge 1$ for $x\in E$.

In order to prove the inequality $\bar{D}^sF(x)\le 1$ for $x\in E$, let $x\in I=[u,v]$, we might as well assume that $u,v\in E$, otherwise we can appropriately reduce I and this will not reduce $F(I)/|I|^s$, therefore $|I|=v-u=\Sigma_{i=n}^{\infty}\alpha_i 5^{-i}$ where $\alpha_n=0,2,$ or 4, $\alpha_i=0,\pm 2,$ or $\pm\ 4$ for all $i\ge n+1$, and

$$F(I) = F(v) - F(u) = \frac{1}{2}\Sigma_{i=n}^{\infty}\frac{\alpha_i}{3^i}\,.$$

We shall only show that

$$\left(\sum_{i=n}^{\infty}\frac{\alpha_i}{5^i}\right)^s \geq \frac{1}{2}\sum_{i=n}^{\infty}\frac{\alpha_i}{3^i}\ . \tag{1}$$

By the fact that the continuity of the power function x^s, it suffices to show that

$$\left(\sum_{i=n}^{n+p}\frac{\alpha_i}{5^i}\right)^s \geq \frac{1}{2}\sum_{i=n}^{n+p}\frac{\alpha_i}{3^i} \tag{2}$$

holds for all nonnegative integer p and $\Sigma_{i=n}^{n+p}\alpha_i 5^{-i} \geq 0$. We shall prove (2) by induction.

In the first place, let $p=0$, it is obvious that (2) holds when $\alpha_n = 0$; moreover, if $\alpha_n = 2$ or $\alpha_n = 4$, the inequality (2) holds, respectively, by the fact that

$$\left(\frac{2}{5^n}\right)^s = \frac{2^s}{3^n} > \frac{1}{3^n} = \frac{1}{2}\cdot\frac{2}{3^n}\ , \text{ and } \left(\frac{4}{5^n}\right)^s = \frac{4^s}{3^n} > \frac{2.57}{3^n} > \frac{1}{2}\cdot\frac{4}{3^n}\ . \tag{3}$$

In the next place, suppose that the inequality (2) holds for $p-1$. To obtain the inequality (2) for p, we will only to show that

$$\left(\sum_{i=n}^{n+p}\frac{\alpha_i}{5^i}\right)^s \geq \frac{1}{2}\cdot\frac{\alpha_n}{3^n} + \left(\sum_{i=n+1}^{n+p}\frac{\alpha_i}{5^i}\right)^s\ . \tag{4}$$

When $\alpha_n = 0$, notice that $\Sigma_{i=n+1}^{n+p}\alpha_i 5^{-i} = \Sigma_{i=n}^{n+p}\alpha_i 5^{-i} \geq 0$, the inequality (2) holds by the hypothesis of the induction. Moreover, when $\alpha_n = 2$ or $\alpha_n = 4$, if $\Sigma_{i=n+1}^{n+p}\alpha_i 5^{-i} \geq 0$, consider the function

$$g(y) = \left(\frac{\alpha_n}{5^n} + y\right)^s - \frac{1}{2}\cdot\frac{\alpha_n}{3^n} - y^s\ ,\quad y \in [0, 5^{-n}]\ .$$

Since

$$g'(y) = s\left(\frac{\alpha_n}{5^n} + y\right)^{s-1} - sy^{s-1} = sy^{s-1}\left[\left(\frac{y}{\alpha_n 5^{-n} + y}\right)^{1-s} - 1\right] < 0\ ,\quad y \in [0, 5^{-n}]\ ,$$

we see that $g(y)$ is decreasing on $[0, 5^{-n}]$, notice that $\Sigma_{i=n+1}^{n+p}\alpha_i 5^{-i} \leq 5^{-n}$, we have

$$\begin{aligned} g(\Sigma_{i=n+1}^{n+p}\alpha_i 5^{-i}) \geq g(5^{-n}) &= \left(\frac{\alpha_n}{5^n} + \frac{1}{5^n}\right)^s - \frac{1}{2}\cdot\frac{\alpha_n}{3^n} - \left(\frac{1}{5^n}\right)^s \\ &= \frac{1}{2}\cdot\frac{1}{3^n}\left[2(\alpha_n + 1)^s - (\alpha_n + 2)\right] \end{aligned}$$

$$\geq \left\{ \begin{array}{ll} \frac{1}{2}\cdot\frac{1}{3^n}\left[2\cdot 2.116-(2+2)\right], & \text{when } \alpha_n = 2, \\ \frac{1}{2}\cdot\frac{1}{3^n}\left[2\cdot 3-(4+2)\right], & \text{when } \alpha_n = 4, \end{array} \right\} \geq 0$$

and by (3), we have

$$g(0) = \left(\frac{\alpha_n}{5^n}\right)^s - \frac{1}{2}\cdot\frac{\alpha_n}{3^n} = \frac{1}{3^n}\left(\alpha_n{}^s - \frac{1}{2}\alpha_n\right) > 0 ,$$

it follows that $g(y) \geq 0$ for all $y \in [0, 5^{-n}]$, and (4) holds whenever $\Sigma_{i=n+1}^{n+p}\alpha_i 5^{-i} \geq 0$. If $\Sigma_{i=n+1}^{n+p}\alpha_i 5^{-i} \leq 0$, in the same way, (4) can be easily proved by considering the function

$$g(y) = \left(\frac{\alpha_n}{5^n} - y\right)^s - \frac{1}{2}\cdot\frac{\alpha_n}{3^n} + y^s , \quad y \in [0, 5^{-n}] .$$

Consequently, the inequality (2) holds for all p and we get $\bar{D}^s F(x) \leq 1$ for $x \in E$.

References

1. Chung, K.L.: A course in probability theory, 2nd edn. Academic Press, New York (1974)
2. Falconer, K.J.: The geometry of fractal sets. Cambridge University Press, Cambridge (1985)
3. Lin, Y.J., Lu, S.P.: Equivalent definitions of Hausdorff measures. C. R. Acad. Sci. Paris t.329 Série I, 947–951 (1999)
4. Lu, S.P.: Hausdorff calculus and some of its applications. Northeast Math. J. 14(4), 479–488 (1998)
5. Lu, S.P.: On rectifiable curves in the plane. Acta. Math. Sci. 17(2), 327–332 (2001)
6. Lu, S.P., Lin, Y.J.: Decomposition of variational measure and the arc-length of a curve in R^n. Real Analysis Exchange 27(2), 673–682 (2001/2002)
7. Saks, S.: Theory of the integral, 2nd English edn., Warsaw (1937)
8. Lu, S.P., Lin, Y.J.: An integral expression of Hausdorff measures in terms of s-densities. C. R. Acad. Sci. Paris t.330 Série I, 11–13 (2000)

A New Artificial Glowworm Swarm Optimization Algorithm Based on Chaos Method

Jun-li Zhang[1], Guo Zhou[2], and Yong-quan Zhou[1]

[1] College of Mathematics and Computer Science,
Guangxi University for Nationalities, Nanning 530006, China
e-mail: yongquanzhou@126.com
[2] Institute of Computing Technology, Chinese Academy of Sciences,
Beijing, 100081, China
e-mail: zhouguo@ict.ac.cn

Abstract. Artificial glowworm swarm optimization (GSO) algorithm can effectively capture all local maxima of the multi-modal function, but it exist some shortcomings for searching the global optimal solution, such as the slow convergence speed, easily falling into the local optimum value, the low computational accuracy and success rate of convergence. According to the chaotic motion with randomness, ergodicity and intrinsic regularity, this paper proposes an improved artificial GSO algorithm based on the chaos optimization mechanism, which adopts the chaotic method to locally optimize the better points that are searched by GSO algorithm. Finally, the experimental results based on the six typical functions shows that the improved algorithm has good convergence efficiency, high convergence precision, and better capability of global search and local optimization.

Keywords: Artificial glowworm swarm optimization, chaos method, function optimization.

1 Introduction

In many scientific areas and engineering computation areas, most of problems that people encounter can be attributed to objective optimization problem. Optimization technique based on mathematics is an applied technology for obtaining the optimal solution or satisfactory solution of a variety of engineering problems. With the wide application of electronic computer, optimization technique gets the rapid development and has long been a focus for researchers all the time. However, practical engineering problems have many characteristics such as large-scale, strong constraint, nonlinear, multi- minimal, multi-objective, the difficulty of modeling and so on, so it is difficult to be solved by some traditional gradient-based algorithms that are sensitive to the initial value and the analytical properties of the functions. In recent years, with the rapid development of computational intelligence theory and technology, people has proposed a variety of social bionic evolution algorithms which don't depend on the initial value and the analytical nature of the objective function, including genetic algorithm (GA) that simulated

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 683–693.
springerlink.com

natural selection and genetic mechanisms in the biological world, ant colony algorithm (ACA) that simulated foraging behavior of ant colony, particle swarm optimization (PSO) algorithm that simulated predatory behavior of bird flock and so on. These algorithms are simple and are also implemented easily. In addition, they have strong robustness and easily integrate with other algorithms. For solving many complex optimization problems, they have been demonstrated their excellent performance and great development potential.

In 2009, artificial glowworm swarm optimization (GSO) algorithm that is proposed by Krishnanand K. N. and Ghose, D who are Indian scholars is a relatively new swarm intelligence optimization algorithm [1]. GSO can capture all local maxima of the multi-modal function and is globally convergent, but the local search ability of GSO is poor, which makes it own deficiencies such as slow convergence speed, easily falling into the local optimum value, low computational accuracy and success rate of convergence when GSO is used to solve global optimization problems. However, chaotic motions are random, ergodic and regular, in order to solve the problems as mentioned above, this paper integrates chaos method into GSO, that is to say, chaos method carries out the local search in the satisfactory solution domains that are obtained by GSO, which makes the algorithm get more accurate solution. Finally, the experimental results based on the six typical functions show that the improved algorithm has better capability of global convergence and local optimization. It is a feasible and effective method for solving function optimization problems.

2 The Basic GSO

GSO algorithm that simulated foraging behavior of glowworm swarm in nature is introduced as a new computational intelligence model. The algorithm is based on the following principles: the glowworms communicate with each other by releasing Lucifer in and the environment, each glowworm just gives its response to the surrounding local environment that is determined by its dynamic decision domain.

Consider the n-dimensional unconstrained function optimization problem:

$$\max f(x) \qquad s.t. \quad x \in S \ . \tag{1}$$

If the problem is $\min f(x) \quad s.t. \quad x \in S$, we suppose $g(x) = -f(x)$ and will change it into $\max g(x)$. Where $f:\ S \to R^1$, $x \in R^n$, $S = \prod_{i=1}^{n} [a_i, b_i]$ is the search space and $a_i < b_i$. Assume that the solution of problem (1) is existent, in other words, the global optimum value $\max f(x)$ is existent, and the set of the global optimum points M is non-empty. The mathematical model of GSO is described as follows:

Assume that $x_i(t) = [x_i^{(1)}(t), x_i^{(2)}(t), \cdots, x_i^{(n)}(t)]$ represents the current location of the i -th glowworm, where t is the iteration counter. $x_i^{(j)}(t)$ denote the j -th component of the location of the i -th glowworm at the t -th iteration $f(x)$ is fitness evaluation function, $l_i(t)$ expresses luciferin value of the i -th glowworm at the t -th iteration. All the glowworms start with the same luciferin value l_0, their luciferin values get updated based on the objective fitness values at their initial positions before they start moving, and the luciferin value update equation is expressed as follows:

$$l_i(t) = (1-\rho)l_i(t-1) + \gamma f(x_i(t)), \tag{2}$$

where $\rho \in [0,1]$ is the ratio of luciferin vaporization, $(1-\rho)$ represents the reflection of the cumulative goodness of the path followed by the glowworms in their current luciferin values γ scales the function fitness values. $r_d^i(t)$ represents the dynamic decision domain of the i -th glowworm at the t -th iteration, which is bounded above by a circular sensor range $r_s (0 < r_d^i \le r_s)$. During the course of the movement, $r_d^i(t)$ is updated according to the following equation:

$$r_d^i(t+1) = \min\{r_s, \max\{0, r_d^i(t) + \beta(n_t - |N_i(t)|)\}\}, \tag{3}$$

where β denotes the rate of change of the neighbourhood range, n_t is the neighborhood threshold, which indirectly controls the number of neighbours of each glowworm. $N_i(t)$ is the set of neighbours of the i -th glowworm at the t -th iteration, which consists of those glowworms that have a relatively higher luciferin value and that are located within a dynamic decision domain $r_d^i(t)$. That is to say

$$N_i(t) = \{j : \|x_j(t) - x_i(t)\| < r_d^i(t); l_i(t) < l_j(t)\}, \tag{4}$$

where $\|\vec{x}\|$ is the norm of $\vec{x}$.

When the i -th glowworm is moving, it need decide direction of movement in accordance with luciferin values of glowworms in its set of neighbours. $P_{ij}(t)$ represents probability of movement of the i -th glowworm moving toward the j -th glowworm in its set of neighbours at the t -th iteration, and it is computed on the basis of the following equation:

$$P_{ij}(t) = \frac{l_j(t) - l_i(t)}{\sum_{k \in N_i(t)} l_k(t) - l_i(t)}. \tag{5}$$

According to probability $P_{ij}(t)$ and roulette method, the i -th glowworm selects the j -th glowworm and moves toward it. Suppose that s is step-size of the movement, $x_i(t+1)$ is calculated based on the following equation [1]:

$$x_i(t+1) = x_i(t) + s\left(\frac{x_j(t) - x_i(t)}{\|x_j(t) - x_i(t)\|}\right). \tag{6}$$

3 Chaos Method

A. Definition of Chaos

Chaos is a common nonlinear phenomenon whose action is complex. It looks like a chaotic process of change, in fact it contains the internal law. Chaos optimization method is regarded as a novel optimization technique. In recent years, it has attracted wide attention from academia and engineering, and has been applied in science and engineering practice. According to chaotic variables with randomness, ergodicity and intrinsic regularity, the optimal variables of a given optimization function are linearly mapped to the chaotic variables in the interval [0, 1], and its search process corresponds to the traversal process in the chaotic orbit, which makes the search process with the ability to escape the local optimum value and may eventually obtain the global optimum solution or satisfactory solution.

In this paper, the logistic equation is used to produce chaos serials, which can be expressed by:

$$x_{k+1} = \mu \cdot x_k (1 - x_k) \ , \ 0 \le x_0 \le 1 , \tag{7}$$

where μ is control parameter, x_k is variable, $k = 0,1,2,\cdots$. Although equation (7) is deterministic, it is easy to prove that when $\mu = 4$ and $x_0 \notin \{0, 0.25, 0.5, 0.75, 1\}$, equation (7) is totally in chaos state. That is to say, if we select any n initial values with slight difference, we will obtain n chaotic variables with different trajectories [2]. In order to better describe the algorithm, we take $\mu = 4$ in the following discussion.

B. Implementation of Chaos Method

Let the optimization problem that chaos method deals with be the functional optimization problem $\max f(X)$ and $X = [x_1, x_2, \cdots, x_n]$ be the optimized variable, k denotes chaotic variables iterating mark, K is the maximum number of chaos iterations, $X^{(k)} = [x_1^k, x_2^k, \cdots, x_n^k]$ represents the variable that is searched by chaos method at the k-th iteration, x_j^k denotes the j-th component of $X^{(k)}$, $X^{(0)}$ is initial value. Let X^* be the current optimal variable and $f(X^*)$ be the current optimal value. The procedure of chaos method can be summarized as follows [3]:

Step 1. Initialization: let $k = 0$, $X^{(k)} = X^{(0)}$, $X^* = X^{(0)}$, $f(X^*) = f(X^{(0)})$, according to equation (8), the optimal variables $x_j^k (j = 1, 2, \cdots, n)$ are mapped to the chaotic variables cx_j^k in the interval [0, 1]:

$$cx_j^k = \frac{x_j^k - x_{\min,j}}{x_{\max,j} - x_{\min,j}} \ , \ j = 1, 2, \cdots, n , \tag{8}$$

where $x_{\max,j}$ and $x_{\min,j}$ denote upper bound and lower bound of the j -th-dimensional variable respectively.

Step 2. According to cx_j^k, equation (9) is used to compute the chaotic variables cx_j^{k+1} at the $(k+1)$ -th iteration:

$$cx_j^{k+1} = 4cx_j^k(1-cx_j^k)\,,\ j=1,2,\cdots,n\,. \tag{9}$$

Step 3. According to equation (10), the chaotic variable cx_j^{k+1} is changed into the optimal variable x_j^{k+1}:

$$x_j^{k+1} = x_{\min,j} + cx_j^{k+1}(x_{\max,j} - x_{\min,j})\,,\ j=1,2,\cdots,n\,. \tag{10}$$

Step 4. According to the new optimal variables x_j^{k+1}, $j=1,2,\cdots,n$, calculate the objective function value $f(X^{k+1})$.

Step 5. If $f(X^{k+1}) > f(X^*)$, then X^* and $f(X^*)$ will be updated by X^{k+1} and $f(X^{k+1})$, or else, X^* and $f(X^*)$ will not be updated.

Step 6. If the maximum number of iterations is met, then stop the iteration, X^* is the optimal variable and $f(X^*)$ is the optimal value; or else, let $k=k+1$, go back to Step 2.

4 GSO with Chaotic Local Search (CLS-GSO)

A. Combination Strategy

GSO with chaotic local search (CLS-GSO) is proposed. In CLS-GSO, chaos method as a local search operator is embedded into GSO, that is to say, during the course of each iteration, firstly, GSO implements the global search, then chaos method implements the local search within the given number of steps for the glowworms whose current fitness values are better than the average fitness value, which leads the swarm to the direction of the optimal solution. For this reason, CLS-GSO avoids the weakness of GSO. That is to say, CLS-GSO doesn't easily fall into the local optimum value, which ensures global convergence and local ergodicity of the algorithm. In addition, CLS-GSO has better convergence efficiency and higher precision. It is easier to escape the local optimum value.

B. Implementation Process

Let $S=\prod_{j=1}^{n}[a_j,b_j]$ and $a_j<b_j$ be the search space of the optimization problem, $x_i(t)=[x_i^{(1)}(t),x_i^{(2)}(t),\cdots,x_i^{(n)}(t)]$ represents the current location of the i -th glowworm at the t -th iteration. $f(x)$ is fitness evaluation function, $r_d^i(t)$ represents the dynamic decision domain of the i -th glowworm at the t -th iteration, which is

bounded above by a circular sensor range $r_s (0 < r_d^i \le r_s)$. $l_i(t)$ expresses the luciferin value of the i -th glowworm at the t -th iteration. Suppose that the number of glowworms is N . The procedure of CLS-GSO can be described as follows:

Step 1. Initialization: All the glowworms start with the same luciferin value l_0, that is to say, let $l_i(0) = l_0$. $t = 1$, here, t denotes the mark of GSO iteration. Initialize the location of each glowworm in the search space. Calculate the fitness value of each glowworm. Initialize the current optimal location X^* and the current optimal value $f(X^*)$ according to the fitness values.

Step 2. Implement GSO for all the glowworms:

(1) Update the luciferin value $l_i(t)$ of each glowworm according to (2);

(2) For each glowworm, calculate $N_i(t)$ and $P_{ij}(t)$ according to (4) and (5);

(3) For each glowworm, according to $P_{ij}(t)$ and roulette method, select the j -th glowworm in $N_i(t)$ and move toward it; Let s be step-size of the movement, calculate $x_i(t+1)$ according to (6); if $x_i(t+1) < a_j$, then let $x_i(t+1) = a_j$; if $x_i(t+1) > b_j$, then let $x_i(t+1) = b_j$;

(4) For each glowworm, calculate $r_d^i(t+1)$ according to (3).

Step 3. Calculate the current fitness values $f(x_i(t))$ of all the glowworms and the current average fitness value f_{avg} , find the best glowworm of the current swarm with the best fitness value according to $f(x_i(t))$. If its fitness value is better than $f(X^*)$, then X^* and $f(X^*)$ will be updated by the location and the fitness value of the current best glowworm.

Step 4. Implement chaotic local search: Find the glowworms whose current fitness values are better than f_{avg} , for these glowworms, according to part 3.2, implement chaotic local search in their search space $\prod_{j=1}^{n}[x_i^{(j)}(t) - r_d^i(t) * r_{i1}^j(t), x_i^{(j)}(t) + r_d^i(t) * r_{i2}^j(t)]$, where $r_{i1}^j(t)$ and $r_{i2}^j(t)$ are two random numbers which are produced by the i -th glowworm at the j -th time at the t -th iteration and which are uniformly distributed in the interval [0, 1]. Calculate the current fitness values of these glowworms, find the best glowworm with the best fitness value according to the current fitness values. If its fitness value is better than $f(X^*)$, then X^* and $f(X^*)$ will be updated by the location and the fitness value of the current best glowworm.

Step 5. If the maximum number of iterations is met, then stop the iteration, X^* is the optimal location and $f(X^*)$ is the optimal value; or else, let $t = t+1$, go back to Step 2.

5 Simulation Experiments

A. Experimental Environment

The GSO and the CLS-GSO are coded in MATLAB R2009a and implemented on 2.61GHz CPU machine with 768MB RAM under Windows XP platform. Algorithm parameters of CLS-GSO and GSO are set as follows: the size of the glowworm swarm is fixed to be $N = 50$, the maximum number of iterations is fixed to be $M = 200$, the maximum number of chaos iterations is set to $K = 10$, the ratio of luciferin vaporization is set to $\rho = 0.4$, the scale of the fitness values is set to $\gamma = 0.6$, the rate of change of the neighbourhood range is set to $\beta = 0.08$, the neighbourhood threshold is set to $n_t = 5$, step-size of the movement is set to $s = 0.03$, the initial luciferin value is set to $l_0 = 5$. For all the functions, the circular sensor range and the initial dynamic decision domain of the glowworms are all considered to be uniform and are set to 2.048, 2π ,50, 2, 10, 3 respectively.

B. The Test Function

To evaluate the performance of CLS-GSO, the following six typical functions are selected for test in this paper. The dimensions of all the functions are 2.

(1) De Jong: F_2 $\min f_1(x) = 100(x_1^2 - x_2)^2 + (1 - x_1)^2$

where $-2.048 \le x_1, x_2 \le 2.048$, the optimal value is 0, the best position is (1, 1).

(2) Eggcrate: $\min f_2(x) = x_1^2 + x_2^2 + 25(\sin^2 x_1 + \sin^2 x_2)$

where $-2\pi \le x_1, x_2 \le 2\pi$, the optimal value is 0, the best position is (0, 0).

(3) Bohachevsky 2 : $\min f_3(x) = x_1^2 + 2x_2^2 - 0.3\cos(3\pi x_1)\cos(4\pi x_2) + 0.3$

where $-50 \le x_1, x_2 \le 50$, the optimal value is 0, the best position is (0, 0).

(4) GP-Goldstein-Price (n=2)

$$\min f_4(x) = [1 + (x_1 + x_2 + 1)^2(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2] \times$$
$$[30 + (2x_1 - 3x_2)^2(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2]$$

where $-2 \le x_1, x_2 \le 2$, the optimal value is 3, the best position is (0, -1).

(5) BR-Branin (n=2)

$$\min f_5(x) = (x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6)^2 + 10(1 - \frac{1}{8\pi})\cos x_1 + 10$$

where $-5 \le x_1 \le 10,\ 0 \le x_2 \le 15$, the optimal value is 0.398, the best positions are (-3.142, 12.275), (3.142, 2.275) and (9.425, 2.425).

(6) Six-hump Camel Back Function

$$\min f_6(x) = (4 - 2.1x_1^2 + \frac{x_1^4}{3})x_1^2 + x_1x_2 + (-4 + 4x_2^2)x_2^2,$$

where $-3 \le x_1 \le 3,\ -2 \le x_2 \le 2$, the optimal value is -1.0316, the best positions are (-0.0898, 0.7126) and (0.0898, -0.7126).

C. Simulation Results

To compare CLS-GSO with GSO for optimization capabilities of different functions, take the following test methods: each algorithm is executed independently fifty times. The quality of the solution that is recorded in Table 1 is measured by the maximum function value, the minimum function value and the mean function value out of fifty runs. A run is regarded as a successful run if the optimal value found in that run lies within 3.5% accuracy of the true optimal value of that function, then the minimum number of function evaluations which is used for that run is recorded. The success rate and the average number of effective evaluation that are recorded in Table 2 are computed by the expressions $(N_v/50) \times 100\%$ and $\sum_{i=1}^{N_v} n_i \Big/ N_v$, where N_v represents the number of successful runs out of 50 runs, n_i represents the minimum number of function evaluations which is used for the i -th successful run in Reference [3].

Table 1. Performance comparison of CLS-GSO and GSO for the functions

Functions	Algorithm	Maximum function value	Minimum function value	Mean function value
f_1	ClS-GSO	8.692331644111955e-005	4.832300149727935e-008	1.231525336414898e-005
	GSO	0.069481924769146	4.925366751172589e-007	0.004032977301367
f_2	CLS-GSO	1.970845834088319e-004	7.756711847019891e-007	5.405170175742343e-005
	GSO	9.491759231967500	2.067894686945386e-006	3.052060957702222
f_3	CLS-GSO	9.562394052475831e-005	7.145132524533082e-008	1.690480489850654e-005
	GSO	3.316878967525414e+002	0.449028280143957	69.637174935101157
f_4	CLS-GSO	3.003547934688247	3.000005168972215	3.000768079543359
	GSO	30.003584179725888	3.000016161631907	4.316462849776215
f_5	CLS-GSO	0.397889282098395	0.397887358167974	0.397887755543500
	GSO	0.419160313915976	0.397887374761043	0.398558422206538
f_6	CLS-GSO	-1.031597429075861	-1.031628437195717	-1.031621810175573
	GSO	-1.017941715400425	-1.031627972793364	-1.031336297851895

Table 2. Performance comparison of CLS-GSO and GSO for the functions

Functions	Algorithm	Success rate	Average number of effective evaluation	Standard deviation
f_1	CLS-GSO	100%	2.840000000000000	2.092046889472704e-005
	GSO	98%	44.306122448979593	0.011158177292723
f_2	CLS-GSO	100%	8.039999999999999	4.901008295353951e-005
	GSO	60%	1.203000000000000e+002	4.425144460663793
f_3	CLS-GSO	100%	37.979999999999997	2.103242121873257e-005
	GSO	0	N/A	67.934702139203452
f_4	CLS-GSO	100%	4.180000000000000	7.915828045939991e-004
	GSO	94%	49.382978723404257	5.551762042395289
f_5	CLS-GSO	100%	6.260000000000000	4.819530491829576e-007
	GSO	98%	1.096938775510204e+002	0.003152674520322
f_6	CLS-GSO	100%	1.680000000000000	7.448846883942019e-006
	GSO	100%	29.440000000000001	0.001933155828244

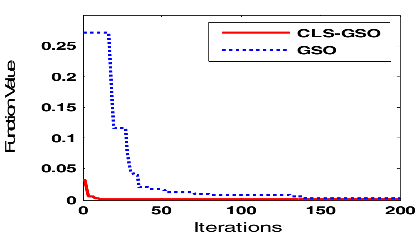


Fig. 1. Evolution curve of the function f_1

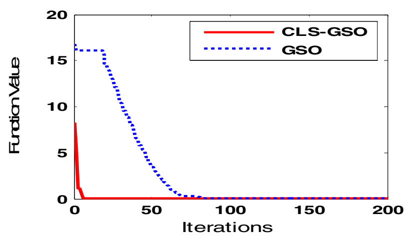


Fig. 2. Evolution curve of the function f_2

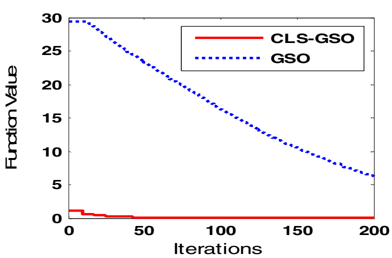


Fig. 3. Evolution curve of the function f_3

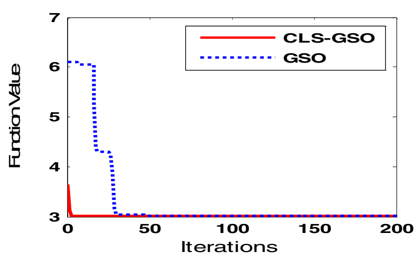


Fig. 4. Evolution curve of the function f_4

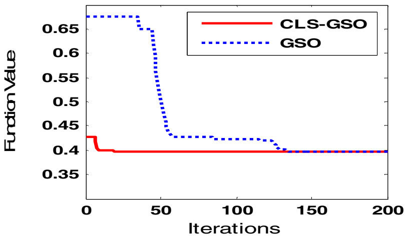


Fig. 5. Evolution curve of the function f_5

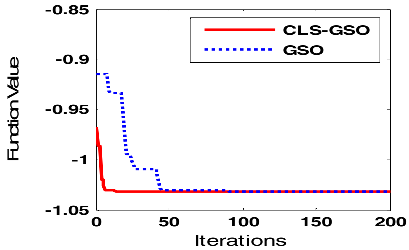


Fig. 6. Evolution curve of the function f_6

D. Analyses of Results

From Table 1, it is observed that CLS-GSO gives a better quality of solutions as compared to GSO, that is to say, for all the functions, the maximum function value, the minimum function value and the mean function value that are obtained by CLS-GSO are better than those that are obtained by GSO from fifty runs. From Table 2, it is observed that CLS-GSO solved all the functions with 100% success rate. However, GSO only solved f_6 with 100% success rate, for the other functions, GSO didn't solve them with 100% success rate. In particular, the success rate of GSO solving f_3 is 0, in other words, GSO could not solve f_3 at all. In addition, for all the functions, the average number of effective evaluation that is used by CLS-GSO is less than of GSO.

From evolution curves of all the functions, it is observed that the convergence speed of CLS-GSO is faster obviously than GSO finding the global optimal solution.

6 Conclusion

In this paper, CLS-GSO for function optimization is presented. It combines chaotic search strategy with GSO, which effectively coordinates the relationship between global search and local search of GSO and which makes GSO not easily fall into the local optimal value. Simulation results show that CLS-GSO outperforms GSO in terms of efficiency, precision, success rate of convergence and reliability.

Acknowledgements. The work was supported by Grants 0832082, 0991086 from Guangxi Science Foundation.

References

1. Krishnanand, K.N., Ghose, D.: Glowworm swarm optimisation: a new method for optimising multi-modal functions, vol. 1(1), pp. 93–119 (2009)
2. Guo, A.-Q., Zhao, Y.-J., Xiong, H., Li, X.: A new class of hybrid particle swarm optimization algorithm. Journal of electronic science and technology of china 5(2), 149–152 (2007)

3. Wang, L., Liu, B.: Particle swarm optimization and scheduling algorithms, pp. 37–39. Tsinghua University Press, Beijing (2008)
4. Zhang, T., Wang, H.-w., Wang, Z.-c.: Mutative scale chaos optimization algorithm and its application. Control and Decision 14(3), 285–288 (1999)
5. Yang, J.-j., Zhou, J.-z.: Particle swarm optimization algorithm based on chaos searching. Computer Engineering and Applications 16, 69–71 (2005)
6. Deep, K., Bansal, J.C.: Mean partical swarm optimization for function optimization. Int.J. Computational Intelligence Studies 1(1), 72–92 (2009)
7. Chen, H., Cui, D.-w., Cui, Y.-A.: Ethnic group evolution algorithm. Journal of Software 20(9), 1–13 (2009) (in Chinese)

Another Metric Model for Trustworthiness of Softwares Based on Partition

Hong-wei Tao and Yi-xiang Chen

Shanghai Key Laboratory of Trustworthy Computing,
East China Normal University, Shanghai, 200062, China
e-mail: tthhww_811@163.com, yxchen@sei.ecnu.edu.cn

Abstract. We once proposed four metric criteria for the multi-dimensional software trustworthiness such as monotonicity, acceleration, sensitivity and substitutivity, and presented two metric models based on theses criteria. Both the substitutivity among critical attributes are 1 and that among non-critical attributes are either 1 or 0 in our two models. In fact the substitutivity between different critical attributes or that between different non-critical attributes should be different. In order to deal with this problem, in this paper we partition critical attributes into several groups. The substitutivity between attributes within the same group are identical, and the substitutivity between attributes belong to different groups are not the same.

Keywords: Trustworthy software, trustworthy metric, multi-dimensional trustworthiness measurement.

1 Introduction

With the increasing demands on software functions, the softwares have been playing an ever-increasing role in our life, the scale and the complexity of the software are getting larger and larger, and environments for software development and running have transited from static closeness to dynamical openness, which leads to a variety of uncertainty factors [1]. Therefore some problems often arise as the softwares work, so how to ensure high trustworthiness of the softwares in the development and operation has become an important research on the theory and technology of software [2]. One of the core scientific problems in constructing trustworthy softwares is how to measure the trustworthiness of the software [1].

Software trustworthiness, as a new concept, is based on such attributes of software as the accuracy, reliability, safety, timeliness, integrity, availability, predictability, survival, controllability, and many other concepts. There is no common explanation about software trustworthiness up to now, however it is

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 695–705.
springerlink.com

widely considered that software trustworthiness can be characterized by many attributes [3-13]. We will not discuss the specific attributes which influence the software trustworthiness in this paper. We believe that there are many attributes referred to trustworthy attributes here that influence the software trustworthiness, and the trustworthy attributes should contain some or all of the quality attributes, including correctness, reliability, safety and so on. The attributes are not completely orthogonal, for example both quality criterion sets of quality factors correctness and reliability which trustworthy softwares must have contain consistency in McCall model [14]. Therefore substitutivity is likely to happen between attributes. The substitutivity of attributes is to describe the event that we can keep the software trustworthiness through the decreasing one attribute value and increasing the other attribute value for different users. Since different trustworthy attributes have different contributions to trustworthiness, we classified trustworthy attributes into two classes: critical and non-critical attributes [6]. Critical attributes are the attributes that a trustworthy software must have, such as reliability, correctness etc.. If any critical attribute value is less than the threshold decided by software, we think this software is not trustworthy. Because different softwares provide different purposes, some softwares may have other attributes except the critical attributes, such as maintainability, portability and so on. We refer to these attributes as non-critical.

Most methods compute the software trustworthiness or software quality by weight sum. In paper [6], we have proved that the weight sum model is not very suitable from the view of substitutivity, since the attributes in this model can completely substitute each other, which is inconsistent with the actual situation. In order to get a appropriate model, we proposed two models in papers [6] and [7] in turn. Both the substitutivity among critical attributes are 1 and that among non-critical attributes are either 1 or 0 in the models prosed in [6] and [7]. In fact the substitutivity between different critical attributes or between different non-critical attributes should be different. In order to deal with this problem, in this paper we partition critical attributes into several groups. The substitutivity between attributes within the same group are identical, and the substitutivity between attributes belong to different groups are not the same.

The paper is organized as follows. In section 2 we describe metric criteria for the multi-dimensional software trustworthiness proposed in [6]. We introduce the improved metric model for software trustworthiness based on partition in section 3. We give a small example in section 4 and in the last section we make the conclusion.

2 Criteria for Trustworthiness

With the same symbols as in reference [6], we suppose that $y_1, \cdots, y_m$ are critical attributes and $y_{m+1}, \cdots, y_{m+s}$ are non-critical attributes ($m+s=n$), and these values are the degree of these attributes of the software. Let T

be a metric function w.r.t. $y_1, \cdots, y_{m+s}$ for the trustworthiness of software. Inspiring from the research results of mathematical modeling in economics, ecology and the environment (see, [15]), we proposed the following metric criteria for the multi-dimensional software trustworthiness in [6]:

(1) **Monotonicity**
It means that the metric function T is monotonically increase with respect to each y_i. That is, the increment of one attribute leads to the increase of the trustworthiness. Thus, we have

$$\partial T/\partial y_i \geq 0.$$

(2) **Acceleration**
Acceleration describes the changing rate of an attribute. Under the case of the increase of only one attribute y_i and keeping of constant for other attributes y_j, $j \neq i$, the efficiency of using the attribute y_i decreases. That means that

$$\partial^2 T/\partial^2 y_i \leq 0.$$

(3) **Sensitivity**
Sensitivity of T about the i-th attribute is the ratio of trustworthiness measurement percentage increase to the percentage increase of i-th attribute y_i expenditure. This sensitivity is $\frac{\partial T}{\partial y_i}\frac{y_i}{T}$ defined as:

$$\frac{\partial T}{\partial y_i}\frac{y_i}{T} = \lim_{\triangle y_i \to 0} \frac{\frac{\triangle T}{T}}{\frac{\triangle y_i}{y_i}}.$$

The minimal critical attribute is more sensitive to T compared with its relative importance.
(4) **Substitutivity**
Substitutivity of y_i and y_j means that we can change their attribute values and do not change the trustworthy degree of a software. The elasticity of attribute substitution is equal to the percentage variation of the ratio of y_i and y_j divided by the percentage variation of the ratio of dy_i and dy_j:

$$\sigma_{ij} = \frac{d(y_i/y_j)}{d(h_{ij})} \times \frac{h_{ij}}{y_i/y_j},\ 1 \leq i,j \leq m+s,\ i \neq j, \tag{1}$$

where

$$h_{ij} = -\frac{\partial T/\partial y_j}{\partial T/\partial y_i} = \frac{dy_i}{dy_j} \quad 1 \leq i,j \leq m+s,\ i \neq j.$$

We use σ_{ij} to express the difficulty of the substitution between the y_i and y_j attributes. Clearly σ_{ij} satisfies $0 \leq \sigma_{ij} \leq \infty$. The bigger σ_{ij} is, the easier substitution between the y_i and y_j is. The attributes y_i and y_j are completely replaceable at $\sigma_{ij} = \infty$ and they are not replaceable at $\sigma_{ij} = 0$.

3 A Metric Model for Trustworthiness of Softwares Based on Partition

Both the substitutivity between critical attributes are 1 and that between non-critical attributes are either 1 or 0 in the models prosed in [6] and [7]. In fact the substitutivity between different critical attributes or that between different non-critical attributes should be different. In order to deal with this problem, inspired by the two-level constant-elasticity-of-substitution production function presented in [16], in this paper we partition critical attributes into several groups. The substitutivity between attributes within the same group are identical, and the substitutivity between attributes belong to different groups are different. We partition $\{1, 2, \cdots m\}$ into S subsets $\{N_1, N_2, \cdots, N_S\}$ and the correspondingly critical attributes into S groups $\{y^{(1)}, \cdots, y^{(S)}\}$ with $y_i \in y^{(s)}$ if $i \in N_s$. The criterion for partitioning is that the critical attributes with same substitutivity are put in one group. To distinguish the importance between critical and non-critical attributes, we use α and β to denote the proportion of critical and non-critical attributes, respectively. We require that $\alpha + \beta = 1$ and that α always is greater than β, i.e. $\alpha > 0.5 > \beta$. The critical attribute groups are proportioned into $\alpha^{(s)}, 1 \leq s \leq S$ with $\sum\limits_{s=1}^{S} \alpha^{(s)} = 1$. All critical attributes within sth $(1 \leq s \leq S)$ group are proportioned into $\alpha_i^{(s)}, i \in N_s$ which satisfy $\sum\limits_{i \in N_s} \alpha_i^{(s)} = 1$. Similarly, we proportion all non-critical attributes into $\beta_{m+1}, \cdots, \beta_{m+s}$ with $\sum\limits_{i=m+1}^{m+s} \beta_i = 1$. Denote the s with $\min\limits_{1 \leq s \leq S} \{y^{(s)}\}$ by min. For simplicity we set $y^{(\min)} = \min\limits_{1 \leq i \leq m} \{y^{(s)}\}$.

Definition 1. *The Metric Model for Trustworthiness of Softwares based on Partition is*

$$\begin{cases} T = \dfrac{10}{11}(\dfrac{y^{(min)}}{10})^{\epsilon} {y^{(1)}}^{\alpha\alpha^{(1)}} \cdots {y^{(S)}}^{\alpha\alpha^{(S)}} + \dfrac{10}{11} y_{m+1}^{\beta\beta_{m+1}} \cdots y_{m+s}^{\beta\beta_{m+s}}, \\ y^{(s)} = (\sum\limits_{N_s} \alpha_i^{(s)} {y_i}^{-\rho_s})^{-\frac{1}{\rho}_s},\ 1 \leq s \leq S, \end{cases} \tag{2}$$

where $0 \leq \epsilon \leq 1 - \alpha^{(\min)}$ is used to control the influence of the critical attribute group with the minimum value which is called the minimum critical attribute group on the trustworthiness of the software, the bigger ϵ, the greater is the influence. $-1 \leq \rho_s < 0$ $(1 \leq s \leq S)$ is a parameter related to substitutivity between attributes, the bigger ρ_s, the more difficult is the substitutivity between attributes.

Claim (1). T is a monotonically increase function.

Proof. Because

$$\begin{cases} \dfrac{\partial T}{\partial y^{(s)}} = \begin{cases} \dfrac{10}{11}\alpha\alpha^{(s)}(\dfrac{y^{(\min)}}{10})^{\epsilon} y^{(1)^{\alpha\alpha^{(1)}}} \cdots y^{(s)^{\alpha\alpha^{(s)}-1}} \cdots y^{(S)^{\alpha\alpha^{(S)}}} & s \neq min \\ \dfrac{10}{11}(\alpha^{(min)} + \epsilon)(\dfrac{y^{(\min)}}{10})^{\epsilon} y^{(1)^{\alpha\alpha^{(1)}}} \cdots y^{(s)^{\alpha\alpha^{(s)}-1}} \cdots y^{(S)^{\alpha\alpha^{(S)}}} & s = min \end{cases} \\ \dfrac{\partial y^{(s)}}{\partial y_i} = \begin{cases} \alpha_i^{(s)}(\sum\limits_{N_s} \alpha_i^{(s)} {y_i}^{-\rho_s})^{-\frac{1}{\rho_s}-1} {y_i}^{-\rho_s-1} & i \in N_s \\ 0 & i \notin N_s \end{cases} \\ \dfrac{\partial T}{\partial y_i} = \dfrac{10}{11}\beta\beta_i y_{m+1}^{\beta\beta_{m+1}} \cdots y_i^{\beta\beta_i-1} \cdots y_{m+s}^{\beta\beta_{m+s}} \quad m+1 \leq i \leq m+s \end{cases}$$

Then

$$\frac{\partial T}{\partial y_i} \geq 0,\ 1 \leq i \leq m+s.$$

i.e. T is monotonically increasing for each $y_i, 1 \leq i \leq m+s$. □

Claim (2). If $1 \leq y_i \leq 10$ for all i $(1 \leq i \leq m+s)$, then $1 \leq T \leq 10$.

Proof. Because for $1 \leq i \leq m+s,\ 1 \leq s \leq S$

$$1 \leq y_i \leq 10 \ and \ \frac{\partial y^{(s)}}{\partial y_i} \geq 0$$

then we can deserve

$$1 = (\sum_{N_s} \alpha_i^{(s)} 1^{-\rho_s})^{-\frac{1}{\rho}_s} \leq y^{(s)} = (\sum_{N_s} \alpha_i^{(s)} {y_i}^{-\rho_s})^{-\frac{1}{\rho}_s} \leq (\sum_{N_s} \alpha_i^{(s)} 10^{-\rho_s})^{-\frac{1}{\rho}_s} = 10$$

By the definition of T and

$$\frac{\partial T}{\partial y^{(s)}} \geq 0,\ 1 \leq s \leq S$$

we have that

$$\frac{10}{11}(\frac{1}{10})^{\varepsilon} + \frac{10}{11} \leq T \leq \frac{10}{11}10^{\alpha} + \frac{10}{11}10^{\beta}.$$

Substituting $0 \leq \varepsilon \leq 1$ in the above inequality, we obtain

$$1 \leq \frac{10}{11}(\frac{1}{10})^{\varepsilon} + \frac{10}{11} \leq T.$$

Let $z = \dfrac{10}{11}10^{\alpha} + \dfrac{10}{11}10^{\beta}$. Because of $\alpha + \beta = 1$ and $\alpha > 0.5 > \beta$, it follows that

$$\frac{\partial z}{\partial \alpha} = \frac{10\ln 10}{11}(10^{\alpha} - 10^{1-\alpha}) = \frac{10\ln 10}{11}\frac{10^{2\alpha} - 10}{10^{\alpha}} > 0.$$

So, z is monotonically increasing for α. Thus $z \leq 10$, i.e., $T \leq 10$. □

Claim (3). T satisfies the acceleration criterion.

Proof. For $1 \leq i \leq m$

$$\frac{\partial^2 T}{\partial^2 y_i} = \frac{\partial^2 T}{\partial^2 y^{(s)}} \left(\frac{\partial y^{(s)}}{\partial y_i}\right)^2 + \frac{\partial T}{\partial y^{(s)}} \frac{\partial^2 y^{(s)}}{\partial^2 y_i}$$

By taking the derivative, we deserve

$$\begin{cases} \dfrac{\partial^2 T}{\partial^2 y^{(s)}} = \begin{cases} \dfrac{10}{11}\alpha\alpha^{(s)}(\alpha\alpha^{(s)}-1)(\dfrac{y^{(\min)}}{10})^{\epsilon} y^{(1)\alpha\alpha^{(1)}} \cdots y^{(s)\alpha\alpha^{(s)}-2} \cdots y^{(S)\alpha\alpha^{(S)}} \\ s \neq min \\ \dfrac{10}{11}(\alpha^{(min)}+\epsilon)(\alpha^{(min)}+\epsilon-1)(\dfrac{y^{(\min)}}{10})^{\epsilon} y^{(1)\alpha\alpha^{(1)}} \cdots \\ \qquad \cdots y^{(s)\alpha\alpha^{(s)}-2} \cdots y^{(S)\alpha\alpha^{(S)}} \\ s = min \end{cases} \\ \dfrac{\partial^2 y^{(s)}}{\partial^2 y_i} = \begin{cases} -\alpha_i^{(s)}(1+\rho_s)(\sum\limits_{N_s} \alpha_i^{(s)} {y_i}^{-\rho_s})^{-\frac{1}{\rho_s}-2} {y_i}^{-\rho_s-2}(\sum\limits_{N_s, j\neq i} \alpha_j^{(s)} {y_j}^{-\rho_s}) & i \in N_s \\ 0 & i \notin N_s \end{cases} \end{cases}$$

Since $0 \leq \alpha, \alpha^{(s)}, \alpha_i^{(s)}, \alpha^{(min)} + \epsilon \leq 1$ *and* $0 \leq 1 + \rho_s < 1$, from the above expression, we can deserve

$$\frac{\partial^2 T}{\partial^2 y^{(s)}} \leq 0, \quad \frac{\partial^2 y^{(s)}}{\partial^2 y_i} \leq 0,\ 1 \leq i \leq m.$$

Therefore

$$\frac{\partial^2 T}{\partial^2 y_i} \leq 0, \quad 1 \leq i \leq m$$

For $m+1 \leq i \leq m+s$

$$\frac{\partial^2 T}{\partial^2 y_i} = \frac{10}{11}\beta_i(\beta_i - 1) y_{m+1}^{\beta_{m+1}} \cdots y_i^{\beta_i - 2} \cdots y_{m+s}^{\beta_{m+s}} \leq 0.$$

Because

$$0 \leq \beta, \beta_i \leq 1,\ m+1 \leq i \leq m+s.$$

Therefore

$$\frac{\partial^2 T}{\partial^2 y_i} \leq 0,\ m+1 \leq i \leq m+s.$$

□

Claim (4). T is sensitive to all attributes.

Proof. Notice that for $1 \leq i \leq m$

$$\frac{\partial T}{\partial y_i}\frac{y_i}{T} = \frac{\partial T}{\partial y^{(s)}}\frac{\partial y^{(s)}}{\partial y_i}\frac{y_i}{T}$$
$$= \begin{cases} \dfrac{10\alpha\alpha^{(s)}\alpha_i^{(s)}}{11T}(\dfrac{y^{(\min)}}{10})^{\epsilon} {y^{(1)}}^{\alpha\alpha^{(1)}} \cdots {y^{(s)}}^{\alpha\alpha^{(s)}+\rho_s} \cdots {y^{(S)}}^{\alpha\alpha^{(S)}} y_i^{-\rho_s} \\ i \in N_s, s \neq \min, \\ \dfrac{10\alpha(\alpha^{(\min)}+\epsilon)\alpha_i^{(\min)}}{11T}(\dfrac{y^{(\min)}}{10})^{\epsilon} {y^{(1)}}^{\alpha\alpha^{(1)}} \cdots \\ \qquad \cdots {y^{(s)}}^{\alpha\alpha^{(\min)}+\rho_{\min}} \cdots {y^{(S)}}^{\alpha\alpha^{(S)}} y_i^{-\rho_{\min}} \\ i \in N_{\min} \end{cases}$$

and for $m+1 \leq i \leq m+s$

$$\frac{\partial T}{\partial y_i}\frac{y_i}{T} = \frac{10T}{11}\beta\beta_i y_{m+1}^{\beta\beta_{m+1}} \cdots y_i^{\beta\beta_i} \cdots y_{m+s}^{\beta\beta_{m+s}}$$

which means that T is sensitive to all attributes. The minimal critical attribute affects on the whole trustworthy degree more than other attributes by adding of ε. □

Claim (5). T has the substitutivity between attributes.

Proof. We first consider the substitutivity between attributes within the same critical attribute group i.e. $i, j \in N_s,\ 1 \leq s \leq S$

$$\sigma_{ij} = \frac{1}{1+\rho_s}$$

which means that the substitutivity between attributes within the same critical attribute group are identical.

It is easy to get that the substitutivity between critical attribute groups are 1.

For the case of non-critical attributes, we have

$$\sigma_{ij} = 1,\ m+1 \leq i, j \leq m+s,\ i \neq j.$$

Now let us consider the substitutivity between attributes which belong to different non-minimal critical attribute groups i.e. $i \in N_s, j \in N_r, r \neq s, r \neq \min, s \neq \min$. By computation, we have

$$\sigma_{ij} = \frac{c+d}{ac+bd}$$

where

$$\begin{cases} a = 1 + \rho_s - \rho_s \alpha_i^{(s)} [\sum_{N_s} \alpha_i^{(s)} {y_i}^{-\rho_s}]^{-1} {y_i}^{-\rho_s} \\ b = 1 + \rho_r - \rho_r \alpha_j^{(r)} [\sum_{N_r} \alpha_i^{(r)} {y_i}^{-\rho_r}]^{-1} {y_j}^{-\rho_r} \\ c = \alpha^{(r)} \alpha_j^{(r)} (\sum_{N_r} \alpha_i^{(r)} {y_i}^{-\rho_r})^{-1} {y_j}^{-\rho_r} \\ d = \alpha^{(s)} \alpha_i^{(s)} (\sum_{N_s} \alpha_i^{(s)} {y_i}^{-\rho_s})^{-1} {y_i}^{-\rho_s} \end{cases}$$

Similarity, for the substitutivity between attribute which belongs to non-minimal critical attribute group and that in minimal critical attribute group i.e. $i \in N_s, j \in N_r, s = min$, we have

$$\sigma_{ij} = \frac{c + d}{ac + bd}$$

where

$$\begin{cases} a = 1 + \rho_s - \rho_s \alpha_i^{(s)} [\sum_{N_s} \alpha_i^{(s)} {y_i}^{-\rho_s}]^{-1} {y_i}^{-\rho_s} \\ b = 1 + \rho_r - \rho_r \alpha_j^{(r)} [\sum_{N_r} \alpha_i^{(r)} {y_i}^{-\rho_r}]^{-1} {y_j}^{-\rho_r} \\ c = \alpha^{(r)} \alpha_j^{(r)} (\sum_{N_r} \alpha_i^{(r)} {y_i}^{-\rho_r})^{-1} {y_j}^{-\rho_r} \\ d = (\alpha^{(s)} + \epsilon) \alpha_i^{(s)} (\sum_{N_s} \alpha_i^{(s)} {y_i}^{-\rho_s})^{-1} {y_i}^{-\rho_s} \end{cases}$$

About the substitutivity between critical and non-critical attributes, we have the following results. Firstly, for the substitutivity between attribute in non-minimal critical attribute group and non-critical attribute i.e. $i \in N_s, m + 1 \leq j \leq n, s \neq \min$

$$\sigma_{ij} = \frac{c + d}{ac + bd}$$

where

$$\begin{cases} a = 1 + \rho_s - (\alpha\alpha^{(s)} + \rho_s) \alpha_i^{(s)} {y^{(s)}}^{\rho_s} {y_i}^{-\rho_s} \\ b = 1 - \beta\beta_j \\ c = \beta\beta_j y_{m+1}^{\beta\beta_{m+1}} \cdots y_j^{\beta\beta_j} \cdots y_n^{\beta\beta_n} \\ d = \alpha\alpha^{(s)} \alpha_i^{(s)} (\frac{y^{(\min)}}{10})^{\epsilon} {y^{(1)}}^{\alpha\alpha^{(1)}} \cdots {y^{(s)}}^{\alpha\alpha^{(s)} + \rho_s} \cdots {y^{(S)}}^{\alpha\alpha^{(S)}} y_i^{-\rho_s} \end{cases}$$

Secondly, for the substitutivity between attribute in minimal critical attribute group and non-critical attribute i.e. $i \in N_s, m + 1 \leq j \leq n, s = \min$

$$\sigma_{ij} = \frac{c+d}{ac+bd}$$

where

$$\begin{cases} a = 1 + \rho_s - [\alpha(\alpha^{(s)} + \epsilon) + \rho_s]\alpha_i^{(s)} {y^{(s)}}^{\rho_s} {y_i}^{-\rho_s} \\ b = 1 - \beta\beta_j \\ c = \beta\beta_j y_{m+1}^{\beta\beta_{m+1}} \cdots y_j^{\beta\beta_j} \cdots y_n^{\beta\beta_n} \\ d = \alpha(\alpha^{(s)} + \epsilon)(\frac{y^{(\min)}}{10})^{\epsilon} {y^{(1)}}^{\alpha\alpha^{(1)}} \cdots {y^{(s)}}^{\alpha\alpha^{(s)}+\rho_s} \cdots {y^{(S)}}^{\alpha\alpha^{(S)}} y_i^{-\rho_s} \end{cases}$$

To sum up the above arguments, T has the substitutivity between attributes. $\square$

Remark 1. In this model we just partition the critical attributes for the model proposed in [6], in fact we can partition both the critical and non-critical attributes for both the models described in [6, 7]. We will not discuss the detail here for these situations.

4 A Small Simulation

This section gives a small simulation for our model. Suppose the number of the critical attributes and non-critical attributes are 3 and 2 separately, i.e., $m = 3, s = 2$. We use our combinational algorithm to compute the weight in the model. Suppose

$$A_1 = \begin{bmatrix} 1 & 2 & 3 \\ 1/2 & 1 & 3 \\ 1/3 & 1/3 & 1 \end{bmatrix}$$

be the positive reciprocal matrix about critical attributes and

$$A_2 = \begin{bmatrix} 1 & 2 \\ 1/2 & 1 \end{bmatrix}$$

be the positive reciprocal matrix about non-critical attributes. By our combinational algorithm, for A_1 the weight vector

$$(\alpha_1, \alpha_2, \alpha_3) = (0.5278, 0.3325, 0.1396)$$

produced by LLSM is the optimal, and for A_2 the weight vector

$$(\alpha_4, \alpha_5) = (0.6667, 0.3333)$$

produced by CSM is the optimal. We suppose $(\alpha, \beta) = (0.9, 0.1)$. Suppose y_1 and y_2 in the same group in this model and

Table 1. Simulations for our metric models

Sim.	y_1	y_2	y_3	y_4	y_5	ϵ	T_1	T_2	T_3
1	8	8	6	8	8	0.001	6.79	6.65	6.81
	8	8	6	8	8	0.01	6.58	6.43	6.76
2	8	8	7	8	8	0.001	6.90	6.76	6.93
	8	8	7	8	8	0.01	6.69	6.54	6.91
3	8	9	6	8	8	0.001	7.00	6.85	7.02
	8	9	6	8	8	0.01	6.77	6.30	6.99
4	9	8	6	8	8	0.001	7.12	6.97	7.59
	9	8	6	8	8	0.01	6.89	6.75	7.56

$$(\alpha_1^{(1)}, \alpha_2^{(1)}) = (\frac{\alpha\alpha_1}{\alpha\alpha_1 + \alpha\alpha_2}, \frac{\alpha\alpha_2}{\alpha\alpha_1 + \alpha\alpha_2}) = (0.614, 0.386)$$

We set $\rho = 0.5$ and $\rho_1 = -0.5$. Let T_1 and T_2 be the metric functions for trustworthiness of software proposed in [6] and [7] respectively, and let T_3 be the model presented in this paper.

Table 1 is a small simulation for our models with parameters described above. From Table 1, we can find these follows. In the three models the increment of the parameter ε leads the decline of whole trustworthy degree. The increment of these attributes leads the increasing of the trustworthy degree. Among these increments, the minimal critical attribute gives the larger increment relative its importance and the most important critical attribute gives the largest increment. And T_3 is always bigger than the previous two models. Therefore if we want to increase the rank of trustworthiness of a software. Firstly, we would better improve the minimal critical attribute, not only because it is easier to be improved but also it is more sensitive to the software trustworthiness compared with its relative importance. Secondly, we can make the most important attribute better.

5 Conclusion

In this paper, we present a new metric model for the software trustworthiness which satisfies all criteria proposed in [6]. We partition critical attributes into several groups to make the substitutivity between critical attributes that belong to different group different. Compared with the previous our two model [6, 7], this model is better in the view of substitutivity. In fact, we can also partition non-critical attributes into different groups in a similar way to make the substitutivity between non-critical attributes that belong to different group different.

Verifying this model by the real cases is the future work. How to identify and measure the attributes which influence the software trustworthiness and how to distinguish the critical and non-critical attributes are what we will do in the future.

Acknowledgments. This article is supported from the Natural Science Foundation of China (No:90718013), Research and Development Project of High-Technology of China (No: 2007AA01Z189).

References

1. Ke, L., Zhiguang, S., Ji, W., et al.: Overview on major research plan of trustworthy software. Bulletin of National Natural Science Foundation of China 22(3), 145–151 (2008)
2. Huwang, C., Ji, W., Wei, D.: High confidence software engineering technologies. Acta Electronica Sinica 31(12A), 2–7 (2003)
3. Zhi-ming, Z., Shi-long, M.: wei, L., et al.: Complexity of software trustworthiness and its dynamical statistical analysis methods. Science in China Series F: Information Sciences 52 (2009)
4. Zhiming, Z., Shilong, M., Wei, L., et al.: Dynamical characteristics of software trustworthiness and their evolutionary complexity. Science in China Series F: Information Sciences 52(8), 1328–1334 (2009)
5. Shanlin, Y., Shuai, D., Wei, C.: Trustworthy software evaluation using utility based evidence theory. Journal of Computer Research and Development 46(7), 1152–1159 (2009) (in Chinese)
6. Tao, H., Chen, Y.: A metric model for trustworthiness of softwares. In: Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology, vol. 3, pp. 69–72 (2009)
7. Tao, H., Chen, Y.: A new metric model for trustworthiness of softwares. In: Proceedings of the International Conference on Information Science and Applications, vol. 2, pp. 651–658 (2010)
8. Hasselbring, W., Reussner, R.: Toward trustworthy software systems. IEEE Computer 39(4), 91–92 (2006)
9. Kirovski, D., Drinic, M., Potkonjak, M.: Enabling trusted software integrity. ACM SIGOPS Operating Systems Review 36(5), 108–120 (2003)
10. Hongbing, Q., Xiaojie, Z., Maozhong, J.: Research on testing-based software credibility measurement and assessment. In: Proceedings of World Congress on Software Engineering, pp. 59–64 (2009)
11. Jeffrey, V.: Trusted softwarep's holy grail. Software Quality Control Journal 11(1), 9–17 (2003)
12. Bernstein, L.: Trustworthy software systems. SIG SOFT Software Engineering Notes 30(1), 4–5 (2005)
13. Steffen, B., Wilhelm, H., Alexandra, P., et al.: Trustworthy software systems: a discussion of basic concepts and terminology. ACM SIGSOFT Software Engineering Notes archive 31(6), 1–18 (2006)
14. Fenton, N.E., Pfleeger, S.L.: Software metrics: a rigorous and practical approach, 2nd edn. International Thomson Publishing (2003)
15. Hritonenko, N., Tatsenko, Y.: Mathematical Modeling in Economics, Ecology and the Environment. Kluwer Academic Publisher, Dordrecht (1999)
16. Sato, K.: A Two-Level Constant-Elasticity-of-Substitution Production Function. Review of Economic Studies 43, 201–218 (1967)

Multi-random Relation and Its Applications

Xin-xing Wu and Yi-xiang Chen

Shanghai Key Laboratory of Trustworthy Computing,
East China Normal University, Shanghai 200062, China
e-mail: xinxingwu@gmail.com, yxchen@sei.ecnu.edu.cn

Abstract. Relation algebra is used to deal with the semantics of programs, the meaning of a program is given by specifying the input-output function that corresponds to the program. However, these programs are not probabilistic programs, because the relation algebra does not contain specific properties for probabilistic programs. In this paper, we extend usual relation to multi-random relation. Then, we give examples to show that multi-random relation can be used to handle the semantics of probabilistic nondeterministic programs.

Keywords: Relation, multi-random relation, multi-random relation space, probability measure, random variable.

1 Introduction

Relation algebra emerged in the 19th century, which culminated in the algebraic logic of Ernst Schröder. The present-day purely equational form or relation algebra is due to the work of Alfred Tarski and his students in the 1940s [1]. Recently, relation algebra is employed as a mathematical tool to deal with the problems in computer science, such as [2-8]. In [4-8], the authors use relation algebra to deal with the semantics of programs (specifications), and consider the input-output semantics of a program (specification) which is given by a relation on its set of states. However, the relation does not contain specific properties for probabilistic programs, so the programs they have dealt with couldn't be probabilistic programs.

In this paper, we firstly extend usual relation to multi-random relation, and obtain some results on multi-random relation. Then, we use multi-random relation to consider the input-output semantics of probabilistic nondeterministic programs.

This paper is organized as follows. In Section 2, we extend usual relation to multi-random relation, and give some interesting propositions and theorems on multi-random relation. In Section 3, by some examples, we show that

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 707–714.
springerlink.com

multi-random relation can be used to consider the input-output semantics of probabilistic nondeterministic programs. The conclusions follow in section 4.

In the rest of this section we introduce some notations which will be used in the later discussion.

(1) Let S be a set, 2^S denotes the power set of S. $|S|$ denotes the cardinality of S.

(2) Ⓡ denotes the set of real numbers.

(3) N denotes the set of natural numbers.

2 Extension to Multi-random Relation Algebra

In this section, we firstly introduce usual relation, then we extend it to multi-random relation.

2.1 Usual Relation

Definition 2.1 [7]. *Usually a relation R on a set S is a subset of the cartesian product of S. The constant relations are the empty relation ($\emptyset$), the universal relation (V) and the identity relation (I). The operations that can be applied to relations are complementation ($^-$), union ($\cup$), intersection ($\cap$), converse ($\breve{}$), composition or relative product ($;$), relative implication ($\triangleleft$), domain (**dom()**), relation-domain ($^<$), power (^{n}) and reflexive transitive closure (*). These are defined as follows:*

(1) $\emptyset \triangleq \{(s,s^{'})|\mathsf{false}\}$.
(2) $V \triangleq \{(s,s^{'})|\mathsf{true}\}$.
(3) $I \triangleq \{(s,s^{'})|s = s^{'}\}$.
(4) $\bar{R} \triangleq \{(s,s^{'})|\neg(s,s^{'}) \in R\}$.
(5) $Q \cup R \triangleq \{(s,s^{'})|(s,s^{'}) \in Q \vee (s,s^{'}) \in R\}$.
(6) $Q \cap R \triangleq \{(s,s^{'})|(s,s^{'}) \in Q \wedge (s,s^{'}) \in R\}$.
(7) $\breve{R} \triangleq \{(s,s^{'})|(s^{'},s) \in R\}$.
(8) $Q;R \triangleq \{(s,s^{'})|\exists(s^{''} :: (s,s^{''}) \in Q \wedge (s^{''},s^{'}) \in R)\}$.
(9) $Q \triangleleft R \triangleq \overline{Q;\overline{R}}$.
(10) $\mathsf{dom}(R) \triangleq \{s|\exists(s^{'} :: (s,s^{'}) \in R\}$.
(11) $R^< \triangleq R;V \cap I$.
(12) $R^n \triangleq \mathsf{if}\, n = 0 \,\mathsf{then}\, I \,\mathsf{else}\, R^{n-1};R$.
(13) $R^* \triangleq \cup(n : n \geq 0 : R^n)$.

Remark 2.1 [7]. *The precedence of the relational operators from highest to lowest is the following: ($^-$),($\breve{}$),($^<$),(^{n}) and (*) bind equally, followed by ;, then by $\triangleleft$, and finally by $\cap$ and $\cup$.*

Definition 2.2 [7]. *A relation R on a set S is functional (or deterministic) if and only if $\breve{R};R \subseteq I$.*

Remark 2.2 [7]. *A relation R is functional if and only if* $\forall s, s^{'} : \exists t : (t,s) \in R \wedge (t,s^{'}) \in R \Longrightarrow s = s^{'}$.

2.2 Multi-random Relation Algebra

Now, we we will extend usual relation to multi-random relation.

Proposition 2.1. *Let S be a set, $\mathfrak{F} = 2^S$. Then $(S, \mathfrak{F})$ is a measurable space.*

Proof. The proof easily follows by the definition of measurable space. □

Remark 2.3. *$\forall x \in S$, let $xR = \{y|(x,y) \in R\}$. Then $(xR, 2^{xR})$ is a measurable space.*

Proposition 2.2. *Let S be a set and R be a relation on S. $\forall x \in S$, $\mu_{xR}(\cdot)$ is a numerically valued set function with domain 2^{xR}, defined as follows: $\forall A \in 2^{xR}$,*

$$\mu_{xR}(A) = \begin{cases} |A|, & A\, is\, a\, finite\, set; \\ +\infty, & A\, is\, an\, infinte\, set. \end{cases}$$

Then, $(xR, 2^{xR}, \mu_{xR})$ is a measure space.

Proof. By proposition 2.1, the proof follows. □

Definition 2.3. *Let S be a set and $R \subseteq S \times S$. Take a fixed $x \in S$, such that $xR \neq \emptyset$, $\mathfrak{F}_\Delta$ is a Borel field of subsets of $\Delta (\subseteq xR)$, $P_x(\cdot)$ is a numerically valued set function: $\mathfrak{F}_\Delta \mapsto [0,1]$, satisfying the following axioms:*

(i) $P_x(\Delta) = 1$.

(ii) $\forall A \in \mathfrak{F}_\Delta, P_x(A) \geq 0$.

(iii) If $\{A_i\}$ is a countable collection of (pairwise) disjoint in $\mathfrak{F}_\Delta$, then $P_x(\cup_i A_i) = \sum_i P_x(A_i)$.

The quadruple $(\{x\}, \Delta, \mathfrak{F}_\Delta, P_x)_{\Delta \subseteq xR}$ is called a multi-random relation space (generated by x and R), x is called the god of $(\{x\}, \Delta, \mathfrak{F}_\Delta, P_x)_{\Delta \subseteq xR}$, $\Delta \subseteq xR$ is called the sample space, the point in Δ is called the sample point. Furthermore,

(iv) If $\cup_{\{\Delta_\gamma | (\{x\}, \Delta_\gamma, \mathfrak{F}_{\Delta_\gamma}, P_x)_{\Delta_\gamma \subseteq xR} \text{ is a multi-random relation space}\}} \Delta_\gamma = xR$ and $\Delta_{\gamma_1} \cap \Delta_{\gamma_2} = \emptyset, \gamma_1 \neq \gamma_2$.

R is then a multi-random relation on $\{x\}$.

Proposition 2.3. *$\forall x \in S$, if $\exists R$ is a usual relation, $xR \neq \emptyset$. Then, $\exists \Delta \subseteq xR (\Delta \neq \emptyset)$, we can structure a multi-random relation space $(\{x\}, \Delta, \mathfrak{F}_\Delta, P_x)_{\Delta \subseteq xR}$ on Δ.*

Proof. By $xR \neq 0$, we can take $y_0 \in xR$. Then, we can construct multi-random relation space as follows:

$$\Delta = \{y\}, \mathfrak{F}_\Delta \{\{y\}, \emptyset\}, P_x(\{y\}) = 1.$$

So, $(\{x\}, \{y\}, \mathfrak{F}_\Delta, P_x)$ is the desired multi-random relation space. □

Definition 2.4. *(1)For a given multi-random relation R, $\forall x \in S$, $xR \neq \emptyset$. If for each $\Delta \in \{\Delta\,|(\{x\}, \Delta, \mathfrak{F}_\Delta, P_x)_{\Delta \subseteq xR}$ is a multi-random relation space$\}$, $\mathfrak{F}_\Delta$ is a trivial Borel field. Then R is called a trivial-multi-random relation on S. The trivial-multi-random relation is just a usual relation.*

(2)For a given multi-random relation R, $\exists x \in S$ such that $xR \neq \emptyset$. $\exists \Delta \in$ $\sup\{|\Delta|\ |(\{x\}, \Delta, \mathfrak{F}_\Delta, P_x)_{\Delta \subseteq xR}$ is a multi-random relation space$\}$, $\mathfrak{F}_\Delta$ is not a trivial Borel field.. Then R is called a partial-multi-random relation on S.

(3)For a given multi-random relation R, $\forall x \in S$, $xR \neq \emptyset$. If for each $\Delta \in \{\Delta\,|(\{x\}, \Delta, \mathfrak{F}_\Delta, P_x)_{\Delta \subseteq xR}$ is a multi-random relation space$\}$, $|\Delta| = |xR|$ and $\mathfrak{F}_\Delta$ is not a trivial Borel field, Then R is called a total-multi-random relation on S. Specially, if $|xR| \equiv 1$, R degenerates into a trivial-multi-random relation.

Collectively these relations are referred to as the multi-random relations. To avoid confusion, if R is a multi-random relation on S, we denote with R_S^P.

Proposition 2.4. *For any usual relation R, we can transform R to a multi-random relation.*

Proof. By proposition 2.3 and definition 2.3, the proof immediately follows. □

Definition 2.5. $(x, y) \in R^P$ *if and only if* $\exists(\{x\}, \Delta, \mathfrak{F}_\Delta, P_x)_{\Delta \subseteq xR^P}$*, such that* $\forall T \in \{T|\{y\} \subseteq T \in \mathfrak{F}_\Delta\}, P_x(T) > 0$.

Definition 2.6

(1′) $\emptyset^P \triangleq \{(s, s')|\mathit{false} \vee \forall R^P_{\{s\}}, \textit{such that}\{s'\} \subseteq B \in \mathfrak{F}_\Delta, \textit{but} P_s(B) = 0\}$.

(2′) $V^P \triangleq \{(s, s')|\exists R^P_{\{s\}}, \forall T \in \{T|\{s'\} \subseteq T \in \mathfrak{F}_\Delta\}, P_s(T) > 0\}$.

(3′) $I^P \triangleq \{(s, s')|s = s' \wedge \exists R^P_{\{s\}}, \forall T \in \{T|\{s'\} \subseteq T \in \mathfrak{F}_\Delta\}, P_s(T) > 0\}$.

(4′) $\overline{R^P} \triangleq \{(s, s')|\neg(s, s') \in R^P \wedge \exists R^P_{\{s'\}}, \forall T \in \{T|\{s\} \subseteq T \in \mathfrak{F}_\Delta\}, P_{s'}(T) > 0\}$.

(5′) $Q^P \cup R^P \triangleq \{(s, s')|(s, s') \in Q^P \vee (s, s') \in R^P\}$.

(6′) $Q^P \cap R^P \triangleq \{(s, s')|(s, s') \in Q^P \wedge (s, s') \in R^P\}$.

(7′) $\breve{R}^P \triangleq \{(s, s')|(s', s) \in R^P \wedge \exists R^P_{\{s\}}, \forall T \in \{T|\{s'\} \subseteq T \in \mathfrak{F}_\Delta\}, P_s(T) > 0\}$.

(8′) $Q^P; R^P \triangleq \{(s, s')|\exists s'' :: (s, s'') \in Q^P \wedge (s'', s') \in R^P\}$.

(9′) $Q^P \triangleleft R^P \triangleq \overline{Q^P; \overline{R^P}}$.

(10′) $\mathit{dom}(R^P) \triangleq \{s|\exists(s' :: (s, s') \in R^P\}$.

(11′) $(R^P)^< \triangleq R^P; V^P \cap I^P$.

(12′) $(R^P)^n \triangleq \mathit{if}\, n = 0\ \mathit{then}\, I\ \mathit{else}\, (R^P)^{n-1}; R^P$.

(13′) $(R^P)^* \triangleq \cup(n : n \geq 0 : (R^P)^n)$,

here, the precedence of the multi-random relational operators is the same with the usual relation.

Remark 2.4. *By these definitions, we can show that*

(a) A multi-random relation R^P *is almost everywhere functional if and only if* $\forall s, s^{'} : \exists t : (t,s) \in R^P \wedge (t,s^{'}) \in R^P \Longrightarrow s = s^{'}$.

(b) A partial identity is a multi-random relation of the form $\{(s,s')|s \in F \wedge s = s^{'} \wedge \exists R^P_{\{s\}}, \forall T \in \{T|\{s^{'}\} \subseteq T \in \mathfrak{F}_\Delta\}, P_s(T) > 0\}$, *for some* $F \subseteq S$.

(c) $(R^P)^< = \{(s,s')|s \in \mathsf{dom}(R^P) \wedge s = s' \wedge \exists R^P_{\{s\}}, \forall T \in \{T|\{s^{'}\} \subseteq T \in \mathfrak{F}_\Delta\}, P_s(T) > 0\}$. *It is a partial identity.*

(d) $|Q^P \cap R^P| > 0$, *but we may have* $Q^P \cap R^P = \emptyset^P$.

Theorem 2.1. *Let* E^P, Q^P, R^P *be multi-random relations. Then:*

(1) $R^P; \emptyset^P = \emptyset^P; R^P = \emptyset^P$.
(2) $R^P; I^P = I^P; R^P = R^P$.
(3) $V^P; V^P = V^P$.
(4) $P^P; (Q^P \cap R^P) \subseteq P^P; Q^P \cap P^P; R^P)$.
(5) $P^P; (Q^P \cup R^P) = P^P; Q^P \cup P^P; R^P)$.
(6) $P^P \subseteq Q^P \Rightarrow R^P; P^P \subseteq R^P; Q^P$.
(7) $P^P \subseteq Q^P \Rightarrow P^P \lhd R^P \subseteq Q^P \lhd R^P$.
(8) $P^P \subseteq Q^P \Rightarrow R^P \lhd P^P \subseteq R^P \lhd Q^P$.
(9) $I^P \lhd R^P = R^P$.
(10) $P^P \lhd Q^P \cap P^P \lhd R^P = P^P \lhd (Q^P \cap R^P)$.
(11) $P^P \lhd R^P \cap Q^P \lhd R^P = (P^P \cap Q^P) \lhd R^P$.
(12) $P^P; Q^P \lhd R^P = P^P \lhd (Q^P \lhd R^P)$.

Proof. By definition 2.3, the proofs easily follow. □

Theorem 2.2. *Let* $|S| = \aleph_1$, *for each* $x \in S$, $\exists R^P_{\{x\}}$, $(\{x\}, \Delta_x, \mathfrak{F}_{\Delta_x}, P_x)_{\Delta_x \subseteq xR^P}$ *is a multi-random relation space. There exists a unique multi-random relation space* $(S, \prod_{x\in S} \Delta_{\{x\}}, \prod_{x \in S} \mathfrak{F}_{\Delta_{\{x\}}}, P)$, *where* P *satisfy that,* $\forall n \in N$, $S_n = \{x_1, x_2, \cdots, x_n\} \subset S$, $A_x \in \mathfrak{F}_{\Delta_{\{x\}}}$, $x \in S_n$, $P(\prod_{x\in S_n} A_x \times \prod_{x \in S\setminus S_n} \Delta_{\{x\}}) = \prod_{x\in S_n} P_x(A_x)$.

Proof. The proof is similar to the proof of Product Probability Theorem in [9]. □

Proposition 2.5. *Let* $\delta_{xR} = |\{\Delta|(\{x\}, \Delta, \mathfrak{F}^{xR}_\Delta, P_x)_{\Delta \subseteq xR^P}$ *is a multi-random relation space*$\}|$. *Then,* $\delta_{xR} \leqq |xR|$.

Proof. By definition 2.3 and definition 2.4, the proof follows. □

Proposition 2.6. *For any* $K \leqq \aleph_1$, $\exists D \subseteq Ⓡ$, *such that* $|D| = k$.

Proof. By $|Ⓡ| = \aleph_1$, the proof follows. □

Corollary 2.1. *Let* $|S| = \aleph_1$, *take a fixed* $x \in S$, *such that* $R^P_{\{x\}}$. *There exists a unique multi-random relation space* $(\{x\}, \prod_{t\in\delta_{xR}} \Delta_t, \prod_{t \in \delta_{xR}} \mathfrak{F}_{\Delta_t}, P)$, *where* P *satisfy that,* $\forall n \in K(\subseteq N)$ *and* $|K| \leq \delta_{xR}$, $\exists D(D \subseteq Ⓡ)$, $|D| = \delta_{xR}$ $S_n = \{t_1, t_2, \cdots, t_n\} \subset D$, $A_t \in \mathfrak{F}^{xR}_{\Delta_t}$, $t \in S_n$, $P(\prod_{t\in S_n} A_t \times \prod_{t\in\delta_{xR}\setminus S_n} \Delta_t) = \prod_{t\in S_n} P_x(A_t)$.

Proof. This follows immediately from theorem 2.2. □

If $R^P_{\{s\}}$ is a total-multi-random relation, then we have the following corollary:

Corollary 2.2. *Let* $|S| = \aleph_1$, $\forall x \in S$, $(\{x\}, xR, \mathfrak{F}_{R^P_{\{x\}}}, P_x)$ *is a multi-random relation space. There exists a unique multi-random relation space* $(S, \prod_{x\in S} xR, \prod_{x\in S} \mathfrak{F}_{R^P_{\{x\}}}, P)$, *where* P *satisfy that,* $\forall n \in N$, $S_n = \{x_1, x_2, \cdots, x_n\} \subset S$, $A_x \in \mathfrak{F}_{R^P_{\{x\}}}$, $x \in S_n$, $P(\prod_{x\in S_n} A_x \times \prod_{x\in S\setminus S_n} xR) = \prod_{x\in S_n} P_x(A_x)$.

Proof. This follows immediately from theorem 2.2. □

3 Applications

In this section, we will show that how to apply multi-random relation to probabilistic programs.

Probabilistic programs were discussed in [10-11], etc.. The following is the abstract syntax of the probabilistic programming language defined in [10], which will be considered in this paper.

$P ::= ABORT|SKIP|x := e|P \triangleleft \alpha \triangleright P|P\,_r\!\oplus P|P \oplus P|P;P|X|\mu X \cdot P(X)$, here, $P_1\,_r\!\oplus P_2$ is a probabilistic choice, in which P_1 is chosen with probability r and P_2 with 1-r.

By the definition of probabilistic programs, we give two examples, through which we will show that how to use multi-random relation to define the input and output semantics of probabilistic programs.

Example 3.1

var x;

$x :=?;$

$x := (2 \times x) \oplus (x\,_{r_1}\!\oplus x + 1) \oplus (x^2\,_{r_2}\!\oplus x^3).$

From fig. 1, we get the multi-random relation that corresponds to the program in example 3.1, which is as follows

$R^P = \{(x, x')|x' = 2x \vee x' = \xi_1 \vee x' = \xi_2\}.$

Example 3.2

var x;

$x :=?;$

$x := x\,_{r_1}\!\oplus x + 1;$

$x := x\,_{r_1}\!\oplus x + 1;$

$x := x\,_{r_1}\!\oplus x + 1;$

$x := x\,_{r_1}\!\oplus x + 1.$

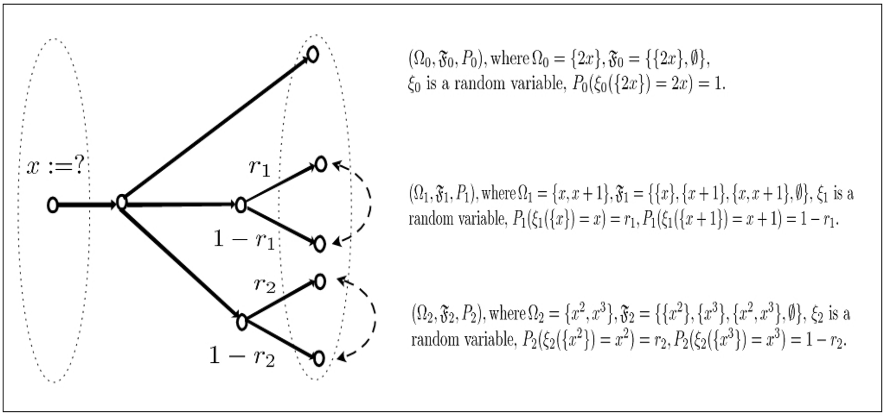


Fig. 1. The relational graph of example 3.1

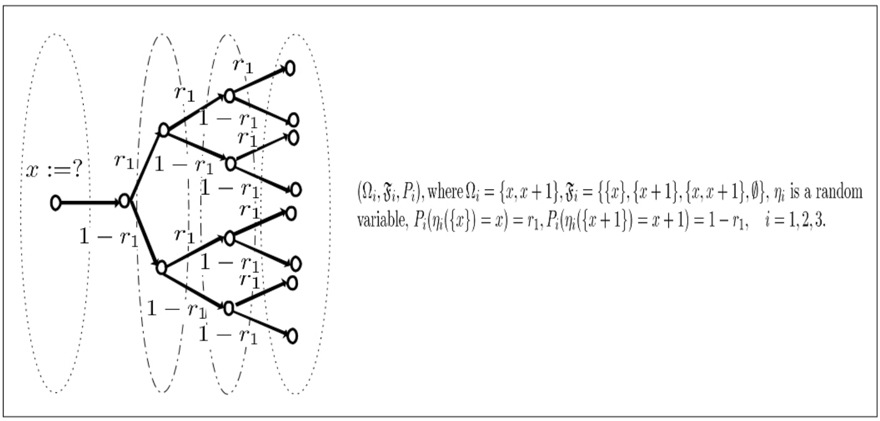


Fig. 2. The relational graph of example 3.2

From fig. 2, we can obtain the multi-random relation that corresponds to the program in example 3.2,

$R^P = \{(x, x')|x' = x + \eta_i\}$, $i = 1, 2, 3$.

So,

$$R^P; R^P; R^P = \{(x, x')|x' = x + \sum_{i=1}^{3} \eta_i\}.$$

4 Conclusion

In this paper, we change the mind's perspective on the relation. We extend the usual relation to multi-random relation, and get some interesting results on multi-random relation. In addition, we give examples to show that multi-random relation can be used to deal with the input-output semantics of probabilistic programs.

Acknowledgements. The research is partially supported by the Natural Science Foundation of China (Grant No. 90718013) , the 973 project (Grant No. 2005CB321904) and the Research and Development Project of High-Technology of China (Grant No. 2007AA01Z189).

References

1. Tarski, A.: On the calculus of relations. J. Symb. Log. 6(3), 73–89 (1941)
2. Brink, C., Kahl, W., Schmidt, G. (eds.): Relational Methods in computer science. Springer, Heidelberg (1997)
3. Schmidt, G., Ströhlein, T.: Relations and graphs. EATCS Monographs in Computer Science. Springer, Berlin (1993)
4. Kahl, W.: Refinement and development of programs from relational specifications. Electronic Notes in Theoretical Computer Science, vol. 44(3), pp. 51–93 (2003)
5. Mili, R., Desharnais, J., Frappier, M., Mili, A.: Semantic distance between specifications. Theoret. Comp. Sci. 247, 257–276 (2000)
6. Jilani, L.L., Desharnais, J., Mili, A.: Defining and applying measures of distance between specifications. IEEE TSE 27(8), 673–703 (2001)
7. Tchier, F., Desharnais, J.: Applying a generalization of a theorem of mills to generalized looping structures. In: Colloquium Science and Engineering for Software Development, organized in the memory of Dr. Harlan D. Mills, and affiliated to the 21st International Conference on Software Engineering, Los Angeles, pp. 31–38 (1999)
8. Tchier, F.: Demonic fixed points. Acta Cybernetica 17(3), 533–555 (2006)
9. Ding, W.D.: Review of measure theory. Anhui People's Publishing House, Hefei (2005) (in Chinese)
10. He, J.F., Seidel, K., McIver, A.K.: Probabilistic models for the guarded command language. Science of Computer Programming 28(2), 171–192 (1997)
11. He, J.F., Sanders, J.W.: Unifying probability. In: Dunne, S., Stoddart, B. (eds.) UTP 2006. LNCS, vol. 4010, pp. 173–199. Springer, Heidelberg (2006)

An Algorithm of Context Rules Generation Based on Virtual Environment

Ming-zhi Chen[1], Chun-yao Xu[1], Guo-rong Cai[3], and Lun Yu[1,2]

[1] College of Physics and Information Engineering,
Fuzhou University, Fuzhou 350108, China
e-mail: donres@fzu.edu.cn
[2] Research Academy of Digital Media, Fuzhou University,
Fuzhou, 350002, China
e-mail: cgifd@126.com
[3] School of Science, Jimei University, Xiamen, 361021, China
e-mail: caiguorong2000@163.com

Abstract. With the widespread application of ubiquitous computing, the context-aware computing is introduced into virtual environment for providing proactive service model in this paper. The key technology of context-aware computing is to predefine the context rules set for inference. The current generation methods of context rules however are usually dependent on manual definition, which could cause lots of problems. Therefore the modified ID3 algorithm is proposed to automatically generate context rules based on virtual environment in this paper. It is to build a context decision tree and then convert this tree into rules set. The experimental result shows that this algorithm has a good performance in effectiveness of generated rules and computational efficiency.

Keywords: Virtual Environment, Context-Aware, Rules Generation, ID3, Decision Tree.

1 Introduction

With the rapid development of network virtual environment, a user might meet multiple services at the same time in virtual environment. For example, the user is likely to both observe the activity state of rival and check-up the use condition of virtual equipments, even to look over the travel path etc., which means that the user need to simultaneously interact with multiple services or multiple systems. The user' attention is distracted because the start-up, configuration for those services and equipments require the user frequently interact with them, so that the user's experience becomes uncomfortable.

Consequently, the concept of ubiquitous computing is introduced into virtual environment to improve the Human-Computer Interaction (HCI) [1]. As yet, there are a lot of researchers to deeply explore the HCI and context-aware computing [2-3]. This computing paradigm is to make the system automatically finding the

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 715–726.
springerlink.com

context information of the user's current situation and surrounding environment, and proactively providing relevant services or resources for the current user, which reduces the frequency of HCI and improves the user's experience [4-5].

The context-aware computing, as the most important branch of ubiquitous computing, was proposed in the early days of 1990s. Schilit B. [6] defined the context environment varying with individuals, and on which the system could improve the interaction of users and computing equipments or other users in his doctoral dissertation in 1994. Ding et al. [7] considered the user's activities may change context, thereby established the prediction model and inference algorithm based on interactive context. Yue et al. [8] investigated how to extract context from interactive environment, and then adjust autologous behaviors. Those research achievements greatly rich the content of context-aware computing and expand the application domain.

The system, solving the problems denoted by context, normally consists of two sections: One is context rules generation, which is to generate context rules and construct context-rules repository. Another is the inference engine to solve the problems. The inference process generally refers to rules match that is to search the rule, which antecedent part is similar to the unsolved problem, in repository.

It is thus clear that the efficient rule-generation is the crucial precondition for the successful inference. How to automatically generate context rules from virtual environment therefore is the research priority in this paper.

2 Problems and Solutions

The inference rules of context-aware system are normally defined by system developer or users by hand. These methods however may bring about some problems as follows [9]. Firstly, owing to system developer's limited cognition to user requirements; it is difficult to define the rules exactly matching user requirements. Secondly, in later period, when those rules need to be modified or improved it is a hard work to express and manage them. Thirdly, there are many redundancies in those manual rules, which may lead to mutual conflict between rules and make rule-matching wrong.

The goal of the context-aware computing is to use the concise and compatible rules for rule matching so that to realize the natural interaction between the users and environments. The complex rules definition makes the interaction more difficult, obviously, which is contradictory to the original intention of context-aware computing.

A new generation method of context rules is proposed in this paper, which is based on virtual environment by means of decision-tree theory. Firstly, we deeply analyze the basic theory of context-aware computing. Afterward, ontology is applied to construct context information modeling in view of networked virtual environment. Secondly, the modified generation algorithm of ID3 decision-tree is carried out. And then it may be automatically transformed into context rules. Finally, the effectiveness of generation algorithm of context rules is verified through experiment.

3 Preliminaries

A. Context awareness concept

Context awareness originated as a term from ubiquitous computing or as so-called pervasive computing which sought to deal with linking changes in the environment with computer systems. The term context-awareness in ubiquitous computing was introduced by Schilit [6]. Context aware devices may also try to make assumptions about the user's current situation. Dey [10] define context as "any information that can be used to characterize the situation of an entity." Context awareness is regarded as an enabling technology for ubiquitous computing systems. Context awareness is used to design innovative user interfaces, and is often used as a part of ubiquitous and wearable computing. It is also beginning to be felt in the internet with the advent of hybrid search engines. Schmidt et al. [11] define human factors and physical environment as two important aspects relating to computer science.

B. Context modeling based on ontology

As a matter of fact, the context modeling is the necessary precondition of context computing. There are two main modeling methods at present: semi-UML language [12] and ontology [13]. The advantage of ontology-based context modeling is that we can describe context in a way with no reference to certain language, which can do formalized analysis for domain knowledge. For example, the First Order Logic could be used to conduct context inference in context-aware computing. Compared with other modeling, the modeling based on ontology has stronger expression power and platform independence [14]. So we adapt ontology to modeling context information.

The relation of ontology and context can be denoted with a triad, $M = \{O, C, R\}$. O is the set of ontology; C is the set of context; R is the relation of ontology and context information.

For example, the *Location(Lisa， Library R01)* indicates that Lisa is in room R01 of Library. Similarly, the *Status (Gate3, Close)* indicates that Gate3 is closed. In virtual environment, the interactions between virtual human and system involve plenty of context information. We use the OWL to construct ontology denoting context information. The relations of core ontology are set up by means of built-in owl: Property [13]. In Fig. 1, we define a set of core context entities: Avatar, Location, Activity, Entity and Time; and then we build inter-relations among those entities.

The owl label is OWL metadata. From the traditional view, it might be an ontology, property or means. In addition, it might be as well the self-defining of certain domain metadata, which describe the information perceived by system when virtual human moving.

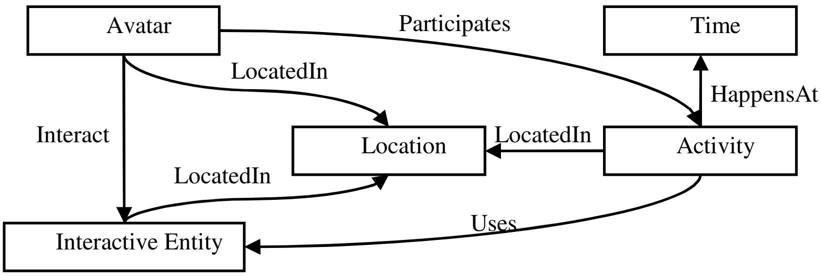


Fig. 1. Inter-relations of ontology-based entities

C. Context information awareness

The context awareness is to transform those factors relevant to current application in computing environment into semantic-explicit and uniform context information.

The context awareness of system starts from the context sensors monitoring and getting environment data. Those context sensors are the interfaces between computing system and context world. They might be physical one or logical one. The logical sensors could be aware of the context of users' interactive habit and history. In this paper, we set the interactive triggers in virtual environment as context sensors. For example, when the virtual human is passing a certain scene and pulls the built-in trigger, system begins to record the context information such as passing time, times and speed etc.. And then that context information is used to construct a context knowledge repository, which is the input resource of context inference rules generation algorithm.

4 Automatic Generation of Context Rules

There are some researchers studying the rules generation algorithm at present [15]. Mitra [16] introduced the fuzzy logic into the primary rule generation algorithm of artificial neural network, which make the algorithm have stronger adaptability. Liu [9] proposed an automatic rule generation algorithm based on rough set theory. He regarded the context-aware system as a decision information system. So he reduced the context information by means of identifiable matrix, and then context rules could be automatically generated.

It is well-known that rough set theory is a mathematic tool to deal with fuzzy and imprecise knowledge, which has a stronger ability of knowledge acquisition. Although the rough set theory is effective for incomplete knowledge, it is quite limited for its fault tolerance and extension ability.

The decision tree learning of artificial intelligence is an inductive learning algorithm based on examples [17]. It has some advantages as follows: Firstly, decision tree could be easier for user understanding. Secondly, the generation efficiency of decision tree is higher than rough set; thus it is more appropriate for large training set. Thirdly, the generation algorithm of decision tree does not require the extra information beside training set. Finally, it could provide more precise for rule generation. Certainly, it might be as well some disadvantages such as unable to exclude the irrelevant attributes with noise and so on, whereas it could be solved via the preprocessing for primary data. Considering the rich context information in virtual environment, the efficiency of rule generation and precision of inference are more emphasized in application. Evidently, the technology of decision tree is more suitable for context rule generation in virtual environment.

The technology of decision tree mainly consists of ID3, C4.5, SLIQ, SPRINT algorithm and so on, in which the most influential algorithm is ID3 presented by R.Quilan in 1986. The others are the variation or modified one of ID3 according to different application. Beside above advantages of decision tree, the reasons we chose ID3 as rule generation algorithm are listed as follows: Firstly, in spite of abundance of context information in virtual environment, by preprocessing ID3 could be able to deal with them. Secondly, the computing efficiency of ID3 is higher than the others, which is in agreement with application requirement of virtual environment. Finally, in order to void over-fitting, to some extent, occurring in building process of ID3 decision tree, we define the indistinguishable relation to modify the classical ID3.

4.1 Context Decision Tree Based on Modified ID3

The ID3 algorithm employs the information gain as choosing standard for tested attributes. It chooses the attribute, which entropy-reducing is most, as division attribute, which is used to split the training examples set and then to build decision tree.

A path from the root to leaf of tree corresponds to a rule. So the whole decision tree corresponds to a set of rules. The interior nodes of tree are the attributes of examples, while the leaves are the prediction values. By this token, the decision tree has a natural advantage for rules generation.

On the basis of analysis of context-aware system's inference mechanism and decision tree theory, we present a generation algorithm of context decision tree based on modified ID3.

Suppose that n different types of context constitute conditional attributes set $A=\{C_1,C_2,\cdots C_n\}$, we define $CV=\{v_1,v_2,\cdots v_n\}$ as a value of those contexts at a certain time, where v_i is the value of context C_i. Let users' next using type of services, namely *ToS* be decision attribute $ToS=\{ToS_1,ToS_2,\cdots ToS_t\}$. And then we define the contexts vector $U=\{CV_1,CV_2,\cdots CV_m,ToS\}$ as the universe of discourse.

In order to avoid over-fitting, we employ the *indistinguishable relation* (see also definition 1) to modify the ID3 in course of tree building.

Definition 1[9]. *Let $IND(\hat{A})$ be the indistinguishable relation, and then $IND(\hat{A})$ is listed as follows.*

$$IND(\hat{A})=\left\{(CV_i,CV_j)\in CV\times CV \,\middle|\, \forall C_k\in\hat{A}, C_k(CV_i)=C_k(CV_j)\vee Class(CV_i)=Class(CV_j)\right\}$$

Where $\hat{A} \subseteq A = \{C_1, C_2, \cdots C_n\}$ *;* $C_k(CV)$ *denotes the value of vector* CV *to attribute* C_k *;* $Class(CV)$ *is the* CV *- belonging class.*

The input and output of algorithm based on modified ID3 are listed as follows.

- Inputs: $U = \{CV_1, CV_2, \cdots CV_m, ToS\}$, $A = \{C_1, C_2, \cdots C_n\}$, $ToS = \{ToS_1, ToS_2, \cdots ToS_t\}$;
- Outputs: (1) ID3 decision tree; (2) context rules.

The basic steps of algorithm are listed as follows.

(1)ID3 (*U,A,ToS*)

{

Create the root node of tree.

If the whole U satisfies $IND(\hat{A})$, then the tree of single node, root labeled $IND(\hat{A})$ *,* is returned.

If the attributes set A is empty, then the single node root is return, which is labeled as the most prevalent ToS 's value in U.

While($|A|! = 0$)

{

Initialize *InfoGain*=0

For *i*=1 to *n* do

//for every attribute, compute information gain, namely InfoGain, and choose the division attribute or attribute set called the BestA.

{

$$Gain(U, C_i) = Entropy(U) - \sum_{v \in CV_i} \frac{|U_v|}{|U|} Entropy(U_v)$$

If $Gain(U, C_i) > InfoGain$ Then $\text{BestA} = C_i$

}

For the every value V_i of BestA

{

- Generate a new branch of root and label it $A = V_i$. Let $U(V_i)$ be the U 's subset which value to A is V_i .
- **If** $U(V_i)$ is empty **Then** generate a leaf under above new branch and label it as the most prevalent ToS 's value in $U(V_i)$, **else** generate a new sub-tree labeled as ID3($U(V_i)$, ToS ,A-BestA) under new branch.

}

}

Return root

}

(2) Transform the ID3 decision into the context rules set.

4.2 Example Analysis for Generation of Context Rules

We illuminate the operation principle of algorithm by means of an example. Firstly, we define the context information such as virtual human location, movement speed, glance time, first time, and user's preference so on. We design a context-aware information service system that can automatically provide user with different type of service information according to current context information. For example, suppose that the user is browsing the art gallery in virtual scene, if the user's representative, virtual human, triggers the sensor built-in somewhere of virtual scene in advance, the system could get those current context information: the movement speed of virtual human is slow; the glance time is relatively long and the user's preference is reading. Therefore, the system introduces the relative background information in text form in order to satisfy the user's personalized interest and special requirement.

We define the context information as a 5-dimensional vector $CV = \{C_1, C_2, C_3, C_4, C_5\}$, where C_1 denotes current location of virtual human, C_2 is movement speed, C_3 is glance time, C_4 denotes whether virtual human is first time to be here, C_5 is user's preference. In order to preferably explain the problem, we advisably simplify the range of value of context attributes. As a result, the movement speed has only three types of value such as high, medium and slow. The value of glance time is long, medium or short. Whether be here first time: yes or no. the user's preference is listening or reading. Accordingly, we define the Table 1 in which a set of context information records and the corresponding type of service are given out.

Table 1. The records of context information and type of services

No	Location(C1)	Movement Speed(C2)	Glance Time(C3)	First Time(C4)	User's Preference(C5)	Type of Service (ToS)
1	Corridor1	High	Long	Yes	Listening	Video
2	ArtGallery1	Slow	Medium	No	Reading	Text
3	ArtGallery2	Medium	Short	No	Listening	Abstract
4	GameArea1	Slow	Medium	Yes	Listening	Video
5	Workshop1	Medium	Medium	No	Reading	Text
6	ScenicArea1	Slow	Long	Yes	Listening	Video
7	GameArea2	Medium	Medium	No	Listening	Audio
8	Corridor2	Medium	Medium	No	Listening	Audio
9	ScenicArea2	High	Medium	No	Reading	Abstract
10	Workshop2	High	Short	No	Reading	Abstract
11	Corridor3	Medium	Long	No	Reading	Text
12	GameArea3	Slow	Medium	No	Reading	Text
13	ArtGallery3	High	Medium	Yes	Listening	Audio
…	…	…	…	…	…	…

Evidently, it is a quite hard work to directly define context rules by those context records in Table 1, which even might some generate redundant and paradoxical rules in spite of the help of Graphical User Interface [18]. The rule-generation algorithm mentioned in this paper does not seek to change the current human-computer interactive model, but make use of a small quantity of interactive record and context information to automatically generate rules. We apply the modified ID3 generation algorithm of decision tree to Table 1, and then can get the decision tree as shown in Fig. 2.

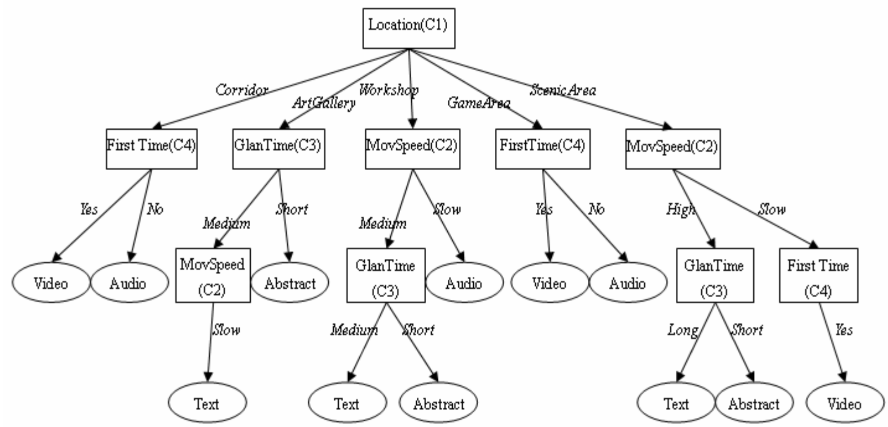


Fig. 2. Context-aware ID3 decision tree

Let the context operation records be antecedent part of rules and the type of service be succedent part. We may get 12 pieces of rules as seen in Table 2.

Table 2. Context-aware rules

No.	Rules			
1	If	(C1□ *Corridor*) ∧ (C4□ *Yes*)	Then	ToS=Video
2	If	(C1□ *Corridor*) ∧ (C4□ *No*)	Then	ToS=Audio
3	If	(C1□ *ArtGallery*) ∧ (C3□ *Medium*) ∧ (C2□ *Slow*)	Then	ToS=Text
4	If	(C1□ *ArtGallery*) ∧ (C3□ *Short*)	Then	ToS=Abstract
5	If	(C1□ *Workshop*) ∧ (C2□ *Medium*) ∧ (C3□ *Medium*)	Then	ToS=Text
6	If	(C1□ *Workshop*) ∧ (C2□ *Medium*) ∧ (C3□ *Short*)	Then	ToS=Abstract
7	If	(C1□ *Workshop*) ∧ (C2□ *Slow*)	Then	ToS=Audio
8	If	(C1□ *GameArea*) ∧ (C4□ *Yes*)	Then	ToS=Video
9	If	(C1□ *GameArea*) ∧ (C4□ *No*)	Then	ToS=Audio
10	If	(C1□ *ScenicArea*) ∧ (C2□ *High*) ∧ (C3□ *Long*)	Then	ToS=Text
11	If	(C1□ *ScenicArea*) ∧ (C2□ *High*) ∧ (C3□ *Short*)	Then	ToS=Abstract
12	If	(C1□ *ScenicArea*) ∧ (C2□ *High*) ∧ (C4□ *Yes*)	Then	ToS=Video

5 Experiment and Analysis

In order to verify the efficiency of rule-generation algorithm, the test data set for training and contrasting is necessary. Thus we need to define rational training and contrasting data set. The experimental steps are given as follows. Firstly, we formulate some context rules and randomly produce an amount of context records, then apply those rules to the latter for rule inference. As a result, we get the proactive service (namely ToS) corresponding to those context records. Secondly, after the redundant and noise data are removed from the above context, we assemble the context record and corresponding ToS as test data set. Thirdly, randomly select a portion to be training data set as the input of generation algorithm of decision tree, while the remainder servers as contrasting data set. Fourthly, the rule-generation algorithm is working. Finally, we apply the antecedent part of rules to match the context records in contrasting data set then get the matching result. Afterwards, compare the matching result with corresponding ToS in contrasting data set, the effectiveness of generated rules could be figured out.

Definition 2 [9]. *EGR (Effectiveness of Generated Rules)*
Let MR be the handmade rules set, GR be the rules set generated by ID3, C be the context records in contrasting set, and $apply(R,c)$ *denotes the ToS which is generated by applying the rules R into context record* $c(\text{where } c \in C)$*, then the EGR can be shown as follows:*

$$EGR = \frac{\left\|\{(c, apply(GR,c))\} \cap \{(c, apply(MR,c))\}\right\|}{\|C\|}. \tag{1}$$

Experiment environment: Operation system: Windows XP Professional with SP2; CPU: Intel Pentium® Dual Core 3.20GHz; EMS memory: Kingston 1.50GB; Main board: Intel 955X; VGA Card: NVIDIA Quadro FX540.

We randomly generate the 1000 pieces of context records including 5 types of context such as location, glance time, movement speed, first time and user's preference, to which we apply the rules listed in table 2 to infer those records-corresponding ToS. Accordingly, a context and its corresponding ToS make up an example. All of examples compose the testing set. Similarly, we also randomly generate 1000 pieces of context records only including 3 types of context.

We test the performance of algorithm from two ways: one is the relation of the effectiveness of generated rules and training set size when the number of context is 3 or 5 as seen in Fig. 3; the other is the relation of the computing time of algorithm and training set size when the number of context is 3 or 5 as seen in Fig. 4.

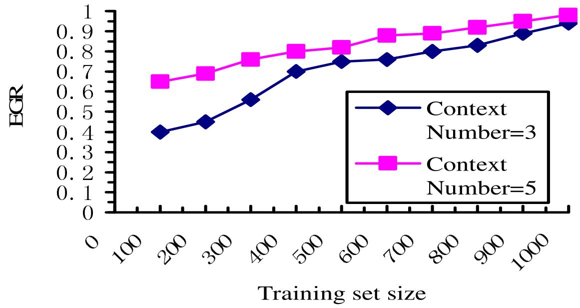


Fig. 3. The relation of the Effectiveness of Generated Rules and training set size

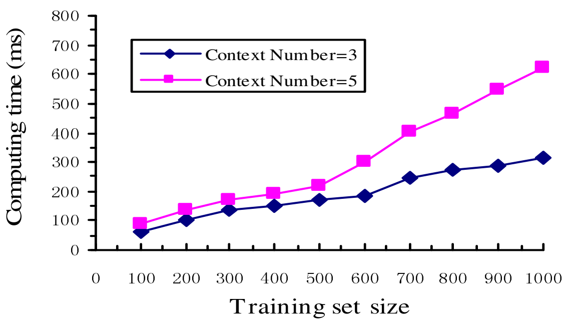


Fig. 4. The relation of the computing time of algorithm and training set size

From Fig 3, we might make the following conclusions. Firstly, the value of EGR is gradually increasing along with the increasing training set size. It finally closes to 1, which shows that the generated rules are approaching the handmade rules. Secondly, for different number of context, the EGR has a different value at the same size of training set. As a general rule, EGR (Context Number=5)> EGR (Context Number =3). The more context numbers are, the more chances there are for ID3 to select best expanding attribute, thus the higher EGR is.

From Fig 4, we might make the following conclusions: Firstly, the computing time of algorithm is near linearly increasing along with the increasing training set size. Secondly, Computing time (Context Number=5)> Computing time (Context Number=3) at the same size of training set. Therefore, on the condition of satisfying the certain EGR requirement, we ought to select those contexts with bigger information gain and limit the number of context to a reasonable extent for fear of unacceptable computing time.

6 Conclusion

In this paper, we, considering the characteristic of virtual environment, mainly explore the context information modeling and context rules generation. One is to utilize the OWL norm to realize the context information modeling by ontologies. The other is to present the modified ID3 algorithm to generate decision tree by using a certain number of context information and users' operation records. And then, the generated decision tree might automatically be transformed into concise and compatible context rules. On that basis, we illuminate the inference principle of context rules to realize proactive service in virtual environment through an example. Finally, the experiment results of EGR show the algorithm mentioned in this paper has a good performance.

Acknowledgements. The work was supported by The Science Research Foundation of Chinese Ministry of Health (No.WKJ2005-2-004) and the Fujian Software Special Project (No. 2006RJ27).

References

1. Guang-you, X., Lin-mi, T., Yuan-chun, S., et al.: Human Computer Interaction for Ubiquitous/Pervasive Computing Mode. Chinese Journal of Computers 30(7), 1041–1053 (2007)
2. Chen, G., Kotz, D.: A Survey of Context-Aware Mobile Computing Research Technical Dartmouth College (2000)
3. Baldauf, M., Dustdar, S., Rosenberg, F.: A survey on context-aware systems. Int. J. Ad. Hoc. Ubiquitous Comput. 2(4), 263–277 (2007)
4. Kim, M., Kim, M.: Fast Service Selection Using rete Network in Dynamic Environment. In: Proceedings of the 2009 Congress on Services - I, pp. 85–92. IEEE Computer Society, Los Alamitos (2009)
5. Yang, S.J.H., Zhang, J., Chen, I.Y.L.: A JESS-enabled context elicitation system for providing context-aware Web services. Expert Syst. Appl. 34(4), 2254–2266 (2008)
6. Schilit, B., Adams, N., Want, R.: Context-Aware Computing Applications. In: Proceedings of the 1994 First Workshop on Mobile Computing Systems and Applications, pp. 85–90. IEEE Computer Society, Los Alamitos (1994)
7. Chun, D., Zhi-yong, F.: Context Prediction Approach Based on Interactive-context. Computer Engineering 35(6), 11–13 (2009)
8. Wei-ning, Y., Yue, W., Guo-ping, W., et al.: Architecture of Intelligent Inter-action Systems Based on Context Awareness. Journal of Computer-Aided Design & Computer Graphics 17(1), 74–79 (2005)
9. Liu, D., Meng, X.W., Chen, J.L., et al.: Algorithms for Rule Generation and Matchmaking in Context-Aware System. Journal of Software 20(10), 2655–2666 (2009)
10. Dey, A.K.: Understanding and Using Context. Personal and Ubiquitous Computing (5), 4–7 (2001)
11. Schmidt, A., Beigl, M., Gellersen, H.: There is more to context than location. Computers & Graphics Journal 23(6), 893–901 (1999)

12. Sheng, Q.Z., Benatallah, B.: Context UML: A UML-Based Modeling Language for Model-Driven Development of Context-Aware Web Services Development. In: International Conference on Mobile Business, pp. 206–212 (2005)
13. Wang, X.H., Zhang, D.Q., Gu, T., et al.: Ontology Based Context Modeling and Reasoning using OWL. In: Proceedings of the Second IEEE Annual Conference on Pervasive Computing and Communications Workshops, vol. 18, pp. 1–5. IEEE Computer Society, Los Alamitos (2004)
14. Gu, T., Pung, H.K., Zhang, D.Q.: A service-oriented middleware for building con-text-aware services. J. Netw. Comput. Appl. 28(1), 1–18 (2005)
15. Jensen, R., Shen, Q.: Fuzzy-rough sets assisted attribute selection. IEEE Transactions on Fuzzy Systems 15(1), 73–89 (2007)
16. Mitra, S., Hayashi, Y.: Neuro-fuzzy rule generation: survey in soft computing frame work, vol. 11(3), pp. 748–768 (2000)
17. Murthy, S.K.: Automatic Construction of Decision Trees from Data: A Multi-Disciplinary Survey. Data Min. Knowl. Discov. 2(4), 345–389 (1998)
18. Nishigaki, K., Yasumoto, K., Shibata, N., et al.: Framework and Rule-Based Language for Facilitating Context-Aware Computing Using Information Appliances. In: Proceedings of the First International Workshop on Services and Infrastructure for the Ubiquitous and Mobile Internet (SIUMI) (ICDCSW 2005), vol. 3, pp. 345–351. IEEE Computer Society, Los Alamitos (2005)

A Classification Method of Grassland and Trees in Remote Sensing Image

Jia-xiang Chen[1], Shui-li Chen[1,2,*], Yun-dong Wu[1,2], and Dan-ping Gui[1]

[1] School of Science
e-mail: sgzx@jmu.edu.cn
[2] Research Center of Image Information Engineering,
Jimei University, Xiamen 361021, China
e-mail: chenjiax@126.com

Abstract. Aimed at the classification of grassland and trees in remote sensing image, according to the image preprocessing of bilateral filter, this method proposed smooth local grassland (trees) similar texture while preserving the edge characteristics between grassland and trees .Then, combining with color image edge detection based on color gradient operator, and dealing with gradient image threshold processing to get binary image, that is , achieve edge detection between grassland and trees, obtain local region rings. Since the topology of segmentation unit is closed, finally making use of mathematical morphology fills with the local closure of the region, obtains tree regions, and ultimately realizes the classification of grassland and trees. The results of experiments show that the algorithm has a good classification of grassland and trees.

Keywords: Bilateral filter, color edge detection, threshold, region filling, classification.

1 Introduction

Grass and trees are green plants, it is difficult for direct classification between grassland and trees in the color space of the remote sensing image. In general, there are different textures between grassland and trees of the Ortho remote sensing images, grassland texture is relatively homogeneous, while tree texture is relatively coarse. In the region of grassland (trees), color (or brightness) and texture features are similar, but the edges between the grassland and trees are more obvious.

Filtering is perhaps the most fundamental operation of image processing and computer vision. In the broadest sense of the term “filtering,” the value of the filtered image at a given location is a function of the values of the input image in a small neighborhood of the same location. Clearly, there is obvious edge information in the image, the usual field operations will inevitably lead to fail at edges, so that the edge of the image blurred. To improve this situation, C. Tomasi and R. Manduchi proposed bilateral filtering method [1], which is non-iterative, local, and simple. It combines gray

* Corresponding author.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 727–734.
springerlink.com

levels or colors based on both their geometric closeness and their photometric similarity, and prefers near values to distant values in both domain and range. The images were smoothed with the template processing, which were not only smoothed locally similar texture, but were largely preserved and strengthened the edges [2]. Therefore, it is useful for grassland and trees on the classification by the use of the image pre-processing of the bilateral filter, which can smooth local grassland (tree) similar texture, and maintain the edge characteristics of grassland and trees.

In this paper, considering texture features of grassland and trees, according to the image preprocessing of bilateral filter, this method proposed smooth local grassland (tree) similar texture while preserving the edge characteristics between grassland and trees .Then, in accordance with color image edge detection based on color gradient operator [3], we obtain gradient image. In order to get the boundary of segmentation unit (local trees region), dealing with gradient image threshold processing gains binary images, and obtains local area rings. Since the topology of segmentation unit is closed, finally making use of mathematical morphology fills with the local closure of the region, obtains tree regions, and ultimately achieves the classification between grassland and trees.

2 Bilateral Filter

Bilateral filter was by C. Tomasi and R. Manduchi, who first put forward in 1998, in contract with traditional Gaussian filters bilateral filtering smooth image while preserving edge [1]. A Gaussian filter applied to image $I_{in}(\vec{x})$ produces an output image $I_{out}(\vec{x})$ defined as follows:

$$I_{out}(\vec{x})=\frac{\sum_{\vec{\varepsilon}\in S_{\vec{x}}} I_{in}(\vec{\varepsilon})\cdot\exp\{-\frac{(\vec{x}-\vec{\varepsilon})^2}{2\delta^2}\}}{\sum_{\vec{\varepsilon}\in S_{\vec{x}}}\exp\{-\frac{(\vec{x}-\vec{\varepsilon})^2}{2\delta^2}\}}, \quad (1)$$

where δ is the standard deviation. This nonlinear filter is widely used to eliminate noise. However, the bilateral filter is a nonlinear filter, the filter coefficients depend on differences in local image pixels, which is combined with two Gaussian filters(one representation is the spatial domain (2D), the other is intensity range). δ_d and δ_r respectively represent geometric spread and photometric spread. Bilateral filter can be expressed as:

$$I_{out}(\vec{x})=\frac{\sum_{\vec{\varepsilon}\in S_{\vec{x}}} I_{in}(\vec{\varepsilon})\cdot\omega_d(\vec{x},\vec{\varepsilon})\cdot\omega_s(\vec{x},\vec{\varepsilon})}{\sum_{\vec{\varepsilon}\in S_{\vec{x}}}\omega_d(\vec{x},\vec{\varepsilon})\cdot\omega_s(\vec{x},\vec{\varepsilon})}, \quad (2)$$

where $S_{\vec{x}}$ is the scope $(2N+1)\times(2N+1)$ of the center $\vec{x}$.

$$\omega_d(\vec{x},\vec{\varepsilon})=\exp\{-\frac{(\vec{x}-\vec{\varepsilon})^2}{2\delta_d^{\,2}}\}, \quad (3)$$

$$\omega_s(\vec{x},\vec{\varepsilon}) = \exp\{-\frac{(I_{in}(\vec{x}) - I_{in}(\vec{\varepsilon}))^2}{2\delta_r^2} \ . \tag{4}$$

Bilateral filter is the product of the nonlinear combination of between $\omega_d(\vec{x},\vec{\varepsilon})$ (spatial proximity factor) and $\omega_s(\vec{x},\vec{\varepsilon})$ (image intensity similarity factor).As the Euclidean distance between the pixel and the center, $\omega_d(\vec{x},\vec{\varepsilon})$ decreases; and $\omega_s(\vec{x},\vec{\varepsilon})$ decreases with the increasing difference between the two pixel brightness values. In the smooth region of the image, neighborhood of pixel brightness values are similar, bilateral filtering changes to Gaussian low-pass filter; whereas, the filter by means of the average brightness values near the edge pixels replace the original in the similar brightness values. Therefore, bilateral filtering smooth images while preserving edges.

Taking advantage of the image preprocessing of bilateral filter, this method proposed smooth local grassland (tree) similar texture while preserving the edge characteristics between grassland and trees. Next, considering how to extract the edges between grassland and trees, here we use color edge detection.

3 Color Edge Detection

Edge detection is an important tool for image segmentation. In this section, we are interested in the issue of computing edges directly in color vector space. The following is one of the various ways in which we can extend in the concept of a gradient to vector functions [3].

Let r , g and b be unit vector along the R , G and B axis of RGB color space and define the vectors[4]

$$u = \frac{\partial R}{\partial x} r + \frac{\partial G}{\partial x} g + \frac{\partial B}{\partial x} b \ , \tag{5}$$

and

$$v = \frac{\partial R}{\partial y} r + \frac{\partial G}{\partial y} g + \frac{\partial B}{\partial y} b \ . \tag{6}$$

The partial derivatives required for implementing Eqs.(5) and (6) can be computed using, for example Sobel operator.

Let the quantities g_{xx} , g_{yy} and g_{xy} be defined in terms of the dot product of these vectors, as follows：

$$g_{xx} = u \cdot u = u^T u = \left|\frac{\partial R}{\partial x}\right|^2 + \left|\frac{\partial G}{\partial x}\right|^2 + \left|\frac{\partial B}{\partial x}\right|^2 \ , \tag{7}$$

$$g_{yy} = v \cdot v = v^T v = \left|\frac{\partial R}{\partial y}\right|^2 + \left|\frac{\partial G}{\partial y}\right|^2 + \left|\frac{\partial B}{\partial y}\right|^2 \ , \tag{8}$$

and

$$g_{xy} = u \cdot v = u^T v = \frac{\partial R}{\partial x}\frac{\partial R}{\partial y} + \frac{\partial G}{\partial x}\frac{\partial G}{\partial y} + \frac{\partial B}{\partial x}\frac{\partial B}{\partial y} . \quad (9)$$

Keeping in mind that R , G and B ,and consequently the g's, are function of x and y .Using this notation, it can be shown[1]that the direction of maximum rate of change of $c(x, y)$ is given by the angle

$$\theta(x, y) = \frac{1}{2}\arctan\left[\frac{2g_{xy}}{(g_{xx} - g_{yy})}\right], \quad (10)$$

and that the value of the rate of change (i.e. the magnitude of the gradient)in the directions given by the elements of $\theta(x, y)$ is given by

$$F_\theta(x, y) = \left\{\frac{1}{2}\left[(g_{xx} + g_{yy}) + (g_{xx} - g_{yy})\cos 2\theta + 2g_{xy}\sin 2\theta\right]\right\}^{1/2} . \quad (11)$$

Note that $\theta(x, y)$ and $F_\theta(x, y)$ are images of the same size as the input image. The elements of $\theta(x, y)$ are simply the angles at each point that the gradient is calculated, and $F_\theta(x, y)$ is the gradient image. Then, we say a pixel at location (x, y) is an edge pixel if $F_\theta(x, y) \geq T$ at the location ,where T is a specified threshold. A threshold image $g(x, y)$ is defined as:

$$g(x, y) = \begin{cases} 1 & F_\theta(x, y) > T \\ 0 & F_\theta(x, y) \leq T \end{cases} . \quad (12)$$

Form Eq.(12) binary image is obtained, pixels labeled 1 correspond to (or objects) boundary points, whereas pixels labeled 0 correspond to the background.

4 Region Filling

From the topology, each segmentation unit (local tree region) is closed, it can be filled with the closed regions on binary image to get some local closure of the regions, which are trees. Next we develop a simple algorithm for region filling based on set dilations, complementation and intersections [4]. In Figure.1, A denotes a set containing a subset whose elements are 8-connected boundary point p inside the boundary ,the objective is to fill the entire region with 1's.If we adopt the convention that all non-boundary（background） points are labeled 0,then we assign a value of 1 to p to begin. The following procedure then fills the region with 1's:

$$X_k = (X_{k-1} \oplus B) \bigcap A^c \qquad k = 1, 2, 3, \ldots , \quad (13)$$

where $X_0 = p$, and B is the symmetric structuring element shown in Fig.1(c).The algorithm terminates at iteration step k if $X_k = X_{k-1}$ .The set union of X_k and A contains the filled set and its boundary. The rest of Figure 1 illustrated further the mechanics of Eq.(13).

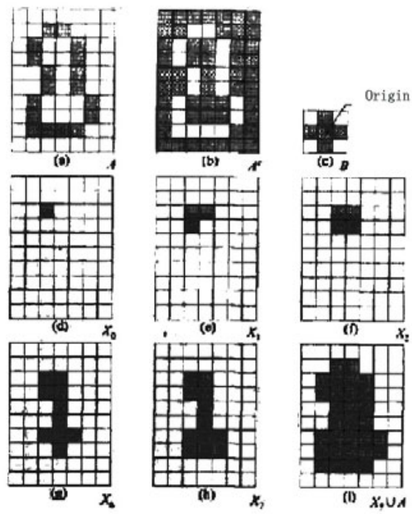


Fig. 1. Region filling：(a) Set A;(b) Complement of A; (c) Structuring elements B; (d) Initial point inside the boundary; (e)- (h) Various steps of Eq.(13); (i) Final result [union of(a)and(h)].

5 Experimental Results and Analysis

In order to test an automatic classification of trees and grassland algorithm in remote sensing image, the images come from low-altitude remote sensing system from Xiamen Passenger Station (image size 512×512). The parameters of the bilateral filter $N = 3$ **,** $\delta_d = 3$ and $\delta_r = 40$,in addition specified threshold $T = 0.12$ **.**

In the process of the experiments, using the bilateral filter not only smooth locally similar texture, but largely preserve and strengthen the edges, in Fig.2 (b), bilateral filter smooth local grassland (tree) similar texture, and maintain the edge characteristics of grassland and trees. Without filtering, original image directly detects edge, gets some discrete points, which are white points in Fig.2 (e); in contract with image preprocessing of bilateral filter and color edge detection, we obtain better closed-loop regions, which are the boundary of all local regions of the trees, shown in Fig.2 (f) boundary points labeled 0. Making use of mathematical morphology fills with the local closure of the regions, which are tree regions shown in Fig. 2 (g), where trees regions are labeled white while grassland regions are black. The dividing lines between trees and grassland are blurred in the image [5], however, grassland and trees can be distinguished. In Fig.2 (h) the red areas are miscarriage by justice subjective analysis, but at last it obtains a good classification between grassland and trees.

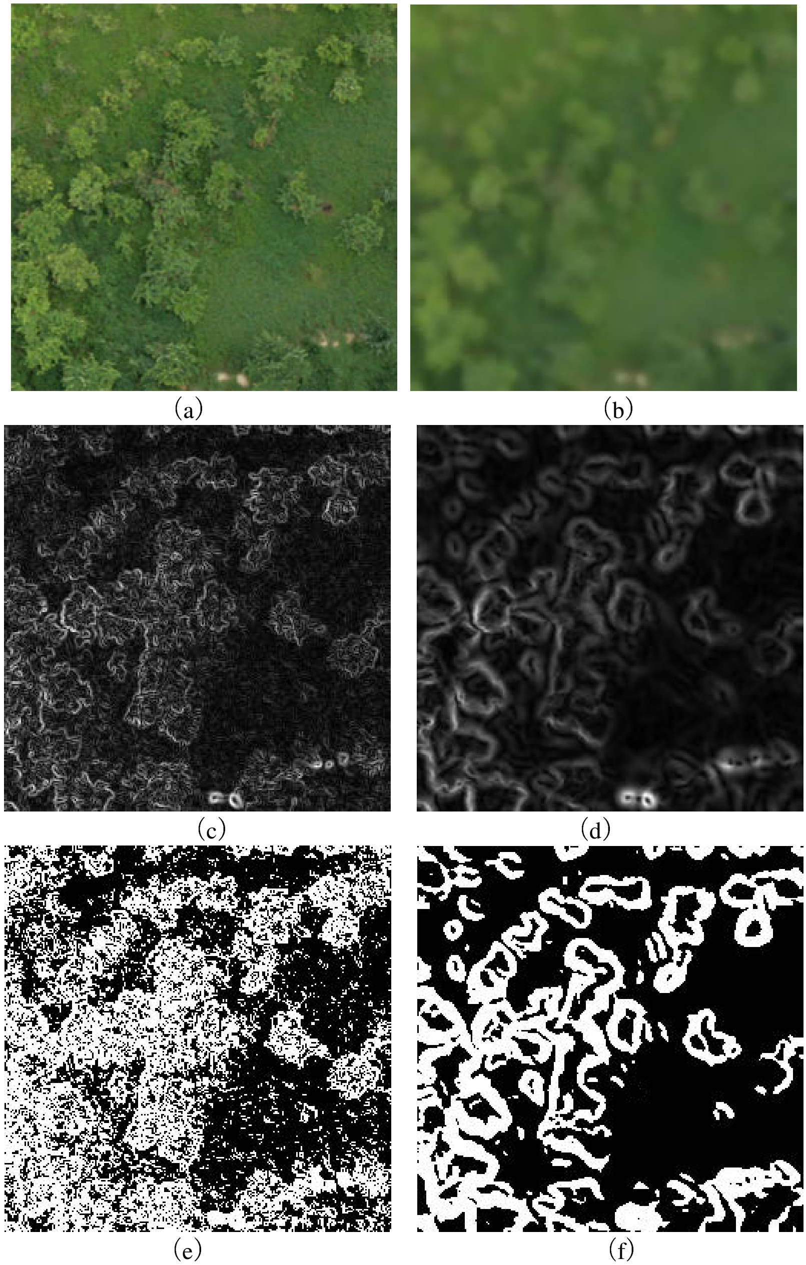


Fig. 2. The procedure of classification between grassland and trees: (a) RGB image;(b) Processed with bilateral filter ; (c) Gradient computed in RGB image(a) vector space; (d) Gradient computed in RGB image(b) vector space; (e) Binary image by image (c) threshold processing ; (f) Binary image by image (d) threshold processing; (g) Image(f) with region filling ; (h) The classification result of the proposed algorithm.

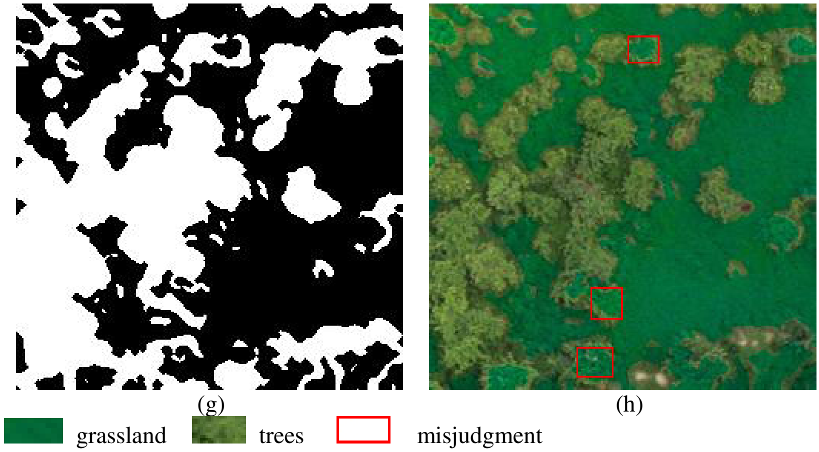


Fig. 2. (*continued*)

6 Conclusion

In this paper we have proposed an automatic classification of trees and grassland algorithm. First of all， according to the image preprocessing of bilateral filter, which not only smooth local grassland (tree) similar texture, but also maintain the edge characteristics of grassland and trees. Next, combining with color image edge detection, obtain binary image. Finally, making use of mathematical morphology, get tree regions, and ultimately achieve the classification of grassland and trees. The automatic classification algorithm has the following distinctive features: the algorithm is unsupervised and no training; the experimental results have shown the preprocessing of the bilateral filter can obtain better closed circuit regions. The next step we will continue to study including making use of collection of fuzzy theory in the classification of grassland and trees [6]; a number of sub-pixels on the edge of the composition need further analysis by sub-segment unit links to extract the closed loop and closed loop non-split unit.

Acknowledgements. This work was supported by the National Natural Science Foundation of Fujian Province (No.A0810014), the key project on Department of Education of Fujian Province (No.JK2009017), and Innovation Fund projects of the younger on Science and Technology Department of Fujian Province (No.2009J05009).

References

1. Tomasi, C., Manduchi, R.: Bilateral Filtering for Gray and Color Images. In: Proceedings of the IEEE International Conference on Computer Vision (1998)
2. Ming, J., Jian-zhong, S.: An adaptive bilateral filtering method for image processing. Opto-Electronic Engineering 31(7), 65–68 (2004)

3. Dizenzo: A Note on the Gradient of a Multi-Image. Computer Vision Graphics and Image Processing, China 33, 1–37 (1986)
4. Gonzalez, R.C.: Digital Image Processing, 2nd edn. Pearson Education, USA (2005)
5. Jensen, J.R.: Introductory Digital Image Processing, 3rd edn. Prentice Hall PTR, Upper Saddle River (2004)
6. Shui-li, C., Jin-gong, L., Xiang-gong, W.: Fuzzy Set Theory and Its Applications. Science Press, China (2005)
7. Overton, K.J., Weymouth, T.E.: A noise reducing preprocessing algorithm. In: Proceedings of IEEE Computer Science Conference on Pattern Recognition and Image Processing, pp. 498–507 (1979)
8. Ying-ying, W.: Bilateral filtering-based multi-edge field of automatic object extraction and description of operator design. Guizhou University, China, Master's thesis (2008)
9. li, M.: Texture Image Analysis. Science Press, Beijing (2009)

P-Set and Its $(f, \bar{f})$-Heredity

Ling Zhang and Xue-fang Ren

School of Mathematics and Computer Science, Longyan University,
Longyan, Fujian 364000, China
e-mail: zl79024@163.com

Abstract. P-set (packet set) is a set pair, which consists of interior P-set (interior packet sets) and exterior P-set (exterior packet sets); by employing the concept of P-set and its structure, the f-heredity set of P-set is presented in this article firstly, then the measurement method and the f-heredity theorems of interior P-set are proposed; similarly the concept of $\bar{f}$-heredity, the measurement method and the $\bar{f}$-heredity theorems of exterior P-set are also proposed. The $(f, \bar{f})$-heredity is one of the important characteristics of P-set.

Keywords: P-set, f-heredity Theorem, $\bar{f}$-heredity Theorem, Heredity Measurement.

1 Introduction

In the year of 2008, by introducing dynamic characteristic into general set (Cantor set) X which has static characteristic, Refs.[1,2] improved general set X, and proposed the concept of P-set (packet set) $(X^{\bar{F}}, X^F)$. P-set is a set pair, which consists of interior P-set (interior packet set) $X^{\bar{F}}$ and exterior P-set (exterior packet set) X^F; P-set $(X^{\bar{F}}, X^F)$ has dynamic characteristics. By employing P-sets theory, some discussions can be abstracted from the fact as follows: suppose $X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\}$ is an apple set, and $\alpha = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ is the attribute set confined in X, where α_1 denotes red color, α_2 denotes sweet taste, α_3 denotes diameter is 6 cm, α_4 denotes weight is 200 g. Due to attribute α_1, set $X^{\bar{F}}_{\alpha_1} = \{x_1, x_3, x_4, x_7, x_8, x_{10}\}$ can be obtained. Obviously that $X^{\bar{F}}_{\alpha_1}$ is the element set with attribute α_1, namely x_1, x_3, x_4, x_7, x_8 and x_{10} are indistinguishable with regard to α_1, the $\mathrm{IND}(X^{\bar{F}}_{\alpha_1})$ exists. Similarly, due to attributes α_1 and α_2, set $X^{\bar{F}}_{\alpha_1,\alpha_2} = \{x_1, x_3, x_4, x_8, x_{10}\}$ can be obtained, and $\mathrm{IND}(X^{\bar{F}}_{\alpha_1,\alpha_2})$ exists; due to attributes α_1, α_2 and α_3, set $X^{\bar{F}}_{\alpha_1,\alpha_2,\alpha_3} = \{x_3, x_4, x_8, x_{10}\}$ can be obtained, and $\mathrm{IND}(X^{\bar{F}}_{\alpha_1,\alpha_2,\alpha_3})$ exists; due to attributes $\alpha_1, \alpha_2, \alpha_3$ and α_4, set $X^{\bar{F}}_{\alpha_1,\alpha_2,\alpha_3,\alpha_4} = \{x_4, x_{10}\}$ can be obtained, and $\mathrm{IND}(X^{\bar{F}}_{\alpha_1,\alpha_2,\alpha_3,\alpha_4})$ exists. The

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 735–743.
springerlink.com

reason of attribute set $\{\alpha_1\}$ becomes $\{\alpha_1, \alpha_2\}$ is the existent of element transfer $f \in F$, and the reason of attribute $\{\alpha_1, \alpha_2\}$ becomes $\{\alpha_1, \alpha_2, \alpha_3\}$ is also the existent of element transfer $f \in F$, and so on. Form where a interesting phenomenon can be inferred that the elements $x_1, x_3, x_4, x_8, x_{10}$ in $X^{\bar{F}}_{\alpha_1}$ are reserved, while element x_7 is deleted. $x_1, x_3, x_4, x_8, x_{10}$ constitute $X^{\bar{F}}_{\alpha_1,\alpha_2} = \{x_1, x_3, x_4, x_8, x_{10}\}$. The elements x_3, x_4, x_8, x_{10} in $X^{\bar{F}}_{\alpha_1,\alpha_2}$ are reserved, x_1 is deleted; x_3, x_4, x_8 and x_{10} constitute $X^{\bar{F}}_{\alpha_1,\alpha_2,\alpha_3} = \{x_3, x_4, x_8, x_{10}\}$. The element x_4 and x_{10} in $X^{\bar{F}}_{\alpha_1,\alpha_2,\alpha_3}$ are reserved, while x_3 and x_8 are deleted; x_4 and x_{10} constitute $X^{\bar{F}}_{\alpha_1,\alpha_2,\alpha_3,\alpha_4} = \{x_4, x_{10}\}$.

The phenomenon above can be introduced in biology, by employing the concept of heredity in biology, a fact can be obtained that if $X^{\bar{F}}_{\alpha_1}$ is great progenitor, $X^{\bar{F}}_{\alpha_1,\alpha_2}$ is progenitor, $X^{\bar{F}}_{\alpha_1,\alpha_2,\alpha_3}$ is parental generation, $X^{\bar{F}}_{\alpha_1,\alpha_2,\alpha_3,\alpha_4}$ is children generation, then x_1, x_3, x_4, x_8 and x_{10} in $X^{\bar{F}}_{\alpha_1}$ are inherited to $X^{\bar{F}}_{\alpha_1,\alpha_2}$. x_1, x_3, x_4, x_8, and x_{10} constitute $X^{\bar{F}}_{\alpha_1,\alpha_2}$. Similarly x_3, x_4, x_8, and x_{10} in $X^{\bar{F}}_{\alpha_1,\alpha_2}$ are inherited to $X^{\bar{F}}_{\alpha_1,\alpha_2,\alpha_3}$, x_3, x_4, x_8 and x_{10} constitute $X^{\bar{F}}_{\alpha_1,\alpha_2,\alpha_3}$. x_4 and x_{10} in $X^{\bar{F}}_{\alpha_1,\alpha_2,\alpha_3}$ are inherited to $X^{\bar{F}}_{\alpha_1,\alpha_2,\alpha_3,\alpha_4}$, x_4 and x_{10} constitute $X^{\bar{F}}_{\alpha_1,\alpha_2,\alpha_3,\alpha_4}$. $x_1, x_3, x_4, x_7, x_8, x_{10}$ present dominance in $X^{\bar{F}}_{\alpha_1}$; x_1, x_3, x_4, x_8 and x_{10} present dominance in $X^{\bar{F}}_{\alpha_1,\alpha_2}$, while x_7 presents recessive in $X^{\bar{F}}_{\alpha_1,\alpha_2}$; x_3, x_4, x_8 and x_{10} present dominance in $X^{\bar{F}}_{\alpha_1,\alpha_2,\alpha_3}$, while x_1 presents recessive in $X^{\bar{F}}_{\alpha_1,\alpha_2,\alpha_3}$, and so on. Form the strict biology option above, a fact can be obtained as follows that interior P-set $X^{\bar{F}}_{\alpha_1,\alpha_2}$ depends on exterior P-set $X^{\bar{F}}_{\alpha_1}$; interior P-set $X^{\bar{F}}_{\alpha_1,\alpha_2,\alpha_3}$ depends on interior P-set $X^{\bar{F}}_{\alpha_1,\alpha_2}$; interior P-set $X^{\bar{F}}_{\alpha_1,\alpha_2,\alpha_3,\alpha_4}$ depends on interior P-set $X^{\bar{F}}_{\alpha_1,\alpha_2,\alpha_3}$. If add attribute α_2 into attribute set $\{\alpha_1\}$, then interior P-set $X^{\bar{F}}_{\alpha_1}$ generates $X^{\bar{F}}_{\alpha_1,\alpha_2}$ hereditarily; if add attribute α_3 into attribute set $\{\alpha_1, \alpha_2\}$, then interior P-set $X^{\bar{F}}_{\alpha_1,\alpha_2}$ generates $X^{\bar{F}}_{\alpha_1,\alpha_2,\alpha_3}$ hereditarily; if add attribute α_4 in attribute set $\{\alpha_1, \alpha_2, \alpha_3\}$, then interior P-set $X^{\bar{F}}_{\alpha_1,\alpha_2,\alpha_3}$ generates $X^{\bar{F}}_{\alpha_1,\alpha_2,\alpha_3,\alpha_4}$hereditarily. From this phenomenon, it is can be concluded that element transfer $f \in F$ changes interior P-set and its structure, leading interior P-set possesses f-heredity. On these grounds, some questions can be proposed as follows: does interior P-set possess f-heredity? In other words, with the change of attribute set, new interior P-set can be obtained from the old one, whether the new interior P-set is the f-heredity of the old one or not? If interior P-set has f-heredity, then what kind of theory or enlightenment about application this fact can offer us? If interior P-set has f-heredity, does it can intercross with biology and give the mathematical interpretation of the biology heredity corresponding to reality? As a mathematical tool, Could interior P-set offer help for the heredity research on biology? Do exterior P-set and interior P-set have same genetic characteristic? By employing the concept of P-set and its structure, this article presents the concepts of P-set and f-heredity set, and proposed the measurement method of interior P-set; then the f-heredity theorem of interior P-set is given. Meanwhile the concept of $\bar{f}$-heredity set of exterior P-set is given, and the measurement method of $\bar{f}$-heredity of P-set and the $\bar{f}$-heredity theorem of exterior P-set are proposed; f-heredity or $\bar{f}$-heredity is a kind of the important characteristics of P-set. For the convenience of discussion and without misunderstanding, moreover keeping the content intact, the P-set and its structure are introduced simply in section 2 as the theory basis and preparation of this article for discussion.

Assumption: $(f, \bar{f})$-heredity is the collective name of f-heredity and $\bar{f}$-heredity.

2 *P*-Set and Its Structure [1, 2]

Assumption: X is the general finite set on U, U is the finite element universe, V is finite attribute universe.

Definition 2.1. *Given the general set* $X = \{x_1, x_2, \cdots, x_m\} \subset U$, $\alpha = \{\alpha_1, \alpha_2, \cdots, \alpha_k\} \subset V$ *is the attribute set of* X. $X^{\bar{F}}$ *is called the interior* P*-set (interior packet set) generated by* X *, or called the interior* P*-set for short, moreover*

$$X^{\bar{F}} = X - X^-, \tag{1}$$

X^-*is called the* $\bar{F}$*-removed element set of* X*, moreover*

$$X^- = \{x|x \in X,\ \bar{f}(x) = u \bar{\in} X,\ \bar{f} \in \bar{F}\}, \tag{2}$$

if the attribute set α^F *of* $X^{\bar{F}}$ *fulfils*

$$\alpha^F = \alpha \cup \{\alpha' | f(\beta) = \alpha' \in \alpha,\ f \in F\}, \tag{3}$$

where $\beta \in V$, $\beta \bar{\in} \alpha$; $f \in F$ *turns* β *into* $f(\beta) = \alpha' \in \alpha$.

Definition 2.2. *Given general set* $X = \{x_1, x_2, \cdots, x_m\} \subset U$, $\alpha = \{\alpha_1, \alpha_2, \cdots,\ \alpha_k\} \subset V$ *is the attribute set of* X, X^F *is called the exterior* P*-set(outer packet set) of* X*, or called the exterior* P*-set for short, moreover*

$$X^F = X \cup X^+, \tag{4}$$

X^+ *is called the* F*-complemented element set of* X*, moreover*

$$X^+ = \{u|u \in U,\ u \bar{\in} X,\ f(u) = x' \in X,\ f \in F\}, \tag{5}$$

if the attribute set $\alpha^{\bar{F}}$ *of* X^F *fulfils*

$$\alpha^{\bar{F}} = \alpha - \{\beta_i | \bar{f}(\alpha_i) = \beta_i \bar{\in} \alpha,\ \bar{f} \in \bar{F}\}, \tag{6}$$

where $\alpha_i \in \alpha$; $\bar{f} \in \bar{F}$ *turns* α_i *into* $\bar{f}(\alpha_i) = \beta_i \bar{\in} \alpha$.

Definition 2.3. *Interior P-set $X^{\bar{F}}$ and exterior P-set X^F constitute a set pair, which is called P-set (packet set) generated by general set, moreover*

$$(X^{\bar{F}}, X^F), \tag{7}$$

$(X^{\bar{F}}, X^F)$ is called P-set for short; the general set X is called the ground set of $(X^{\bar{F}}, X^F)$.

Illustrations of the name P-set

Because of the existent of the element transfer $\bar{f} \in \bar{F}$, in formula (1), the amount of element in the general set X decreases, and X generates $X^{\bar{F}}$, where $X^{\bar{F}}$ is contained in X; similarly, because of the existent of element transfer $f \in F$, in formula (4), the amount of element in the general set X increases, X generates X^F, where X^F contains X; $X^{\bar{F}}$ and X^F are in a state of constant motion. $X^{\bar{F}}$, which is contained in X, and X^F which contains X, constitute P-set $(X^{\bar{F}}, X^F)$ of general set X.

The principle of the generation of P-set

When some elements of a general set X are transferred out from X, or some elements out of X are transferred into X, the general set X generates P-set $(X^{\bar{F}}, X^F)$ which has dynamic characteristic; the existent of $(X^{\bar{F}}, X^F)$ depends on X, and is independent with the amount of elements transferred into or out from X.

3 The f -Heredity and Its Measurement of Interior P-Set

Definition 3.1. *Given general set $X = \{x_1, x_2, \cdots, x_m\} \subset U$, $\alpha = \{\alpha_1, \alpha_2, \cdots, \alpha_k\} \subset V$ is the attribute set of X. $X^{\bar{F}}_{\alpha\cup\{f(\alpha'_i)\}}$ is called the first order f-heredity set of $X^{\bar{F}}$, if there is an attribute $\alpha' \bar{\in} \alpha, f(\alpha') \in \alpha$; similarly $X^{\bar{F}}_{\alpha\cup\{f(\alpha'_i), \cdots, f(\alpha'_j)\}}$ is called the λ order f-heredity set of $X^{\bar{F}}$, and is denoted by $X^{\bar{F}^\lambda}_{(\alpha,f)}$; the parameter λ is called f-genetic order. Where $\lambda = card(\alpha^f)$, $\alpha^f = \{f(\alpha'_i), \cdots, f(\alpha'_j)\}$, and $\lambda \in N^+$.*

Definition 3.2. *$X^{\bar{F}}_\alpha$ is the f-genetic of λ order f-heredity set $X^{\bar{F}^\lambda}_{(\alpha,f)_j}$, if*

$$X^{\bar{F}}_\alpha = \bigcap_{j=1}^{t} X^{\bar{F}^\lambda}_{(\alpha,f)_j}. \tag{8}$$

Definition 3.3. *$GEC(X^{\bar{F}^\lambda}_{(\alpha,f)})$ is called the f-genetic coefficient of λ order f-hereditary set $X^{\bar{F}^\lambda}_{(\alpha,f)}$ with regard to set $X^{\bar{F}}$, moreover*

$$GEC(X^{\bar{F}^\lambda}_{(\alpha,f)}) = GRD(X^{\bar{F}^\lambda}_{(\alpha,f)}) \Big/ GRD(X^{\bar{F}}). \tag{9}$$

Definition 3.4. *$GVD(X^{\bar{F}^{\lambda}}_{(\alpha,f)})$ is called the f-genetic variation degree of λ order f-heredity set $X^{\bar{F}^{\lambda}}_{(\alpha,f)}$ with regard to set $X^{\bar{F}}$, moreover*

$$GVD(X^{\bar{F}^{\lambda}}_{(\alpha,f)}) = 1 - GEC(X^{\bar{F}^{\lambda}}_{(\alpha,f)}). \tag{10}$$

Due to Definitions 3.1-3.4, the following propositions can be obtained.

Proposition 3.1. *f-genetic $X^{\bar{F}}_{\alpha}$ exists in every f-heredity set.*

Proposition 3.2. *Any two λ order f-heredity sets of $X^{\bar{F}}$ fulfil $X^{\bar{F}^{\lambda}}_{(\alpha,f)_i} \cap X^{\bar{F}^{\lambda}}_{(\alpha,f)_j} \neq \emptyset$.*

Propositions 3.1-3.2 is intuitive facts, and the proofs are omitted.

Theorem 3.1. *(The theorem of f-heredity set chain) Suppose $X^{\bar{F}^{\lambda}}_{(\alpha,f)}$ is the λ order f-genetic knowledge of $X^{\bar{F}}$, and $=\lambda_1, \lambda_2, \ldots, \lambda_t$. If $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_t$, then*

$$X^{\bar{F}^{\lambda_t}}_{(\alpha,f)} \subseteq X^{\bar{F}^{\lambda_{t-1}}}_{(\alpha,f)} \subseteq \ldots \subseteq X^{\bar{F}^{\lambda_1}}_{(\alpha,f)}. \tag{11}$$

Proof. Since $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_t$, there are $(\alpha \cup \{f(\alpha'_1)\}) \subseteq (\alpha \cup \{f(\alpha'_1), f(\alpha'_2)\}) \subseteq \cdots \subseteq (\alpha \cup \{f(\alpha'_1), f(\alpha'_2), \cdots, f(\alpha'_\lambda)\})$; $GRD(X^{\bar{F}^{\lambda_t}}_{(\alpha,f)}) \leq GRD(X^{\bar{F}^{\lambda_{t-1}}}_{(\alpha,f)}) \leq \ldots \leq GRD(X^{\bar{F}^{\lambda_1}}_{(\alpha,f)})$ or $card(X^{\bar{F}^{\lambda_t}}_{(\alpha,f)}) \leq card(X^{\bar{F}^{\lambda_{t-1}}}_{(\alpha,f)}) \leq \ldots \leq card(X^{\bar{F}^{\lambda_1}}_{(\alpha,f)}$. So $X^{\bar{F}^{\lambda_t}}_{(\alpha,f)} \subseteq X^{\bar{F}^{\lambda_{t-1}}}_{(\alpha,f)} \subseteq \ldots \subseteq X^{\bar{F}^{\lambda_1}}_{(\alpha,f)}$.

Theorem 3.2. *(The minimum granularity theorem of f-heredity set) Suppose $X^{\bar{F}^{\lambda}}_{(\alpha,f)}$ is the λ order f-heredity set of set $X^{\bar{F}}$. If the attribute set $(\alpha \cup \alpha^f)$ of $X^{\bar{F}^{\lambda}}_{(\alpha,f)}$ fulfils*

$$card(\alpha \cup \alpha^f) = m + \lambda, \tag{12}$$

then $X^{\bar{F}^{\lambda}}_{(\alpha,f)}$ has the minimum granularity $GRD(X^{\bar{F}^{\lambda}}_{(\alpha,f)})$, namely,

$$GRD(X^{\bar{F}^{\lambda}}_{(\alpha,f)}) = \min. \tag{13}$$

Proof. Because $card(\alpha \cup \alpha^f) = m + \lambda$, λ order f-heredity set $X^{\bar{F}^{\lambda}}_{(\alpha,f)}$ has maximum genetic order $\lambda_{\max}=\max\limits_{i=1}^{t}(\lambda_i)$, by employing theorem 3.1, there is

$$card(X^{\bar{F}^{\lambda_t}}_{(\alpha,f)})\Big/card(U) \leq cardX^{\bar{F}^{\lambda_{t-1}}}_{(\alpha,f)})\Big/card(U) \leq \ldots \leq card(X^{\bar{F}^{\lambda_1}}_{(\alpha,f)})\Big/card(U).$$

Then due to the set granularity, there is $GRD(X^{\bar{F}_{\lambda}}_{(\alpha,f)}) = card(X^{\bar{F}_{\lambda_t}}_{(\alpha,f)})\Big/card(U)= \min$.

Theorem 3.3. *(The relation theorem between f-genetic coefficient and f-genetic variation degree) Suppose $GEC(X^{\bar{F}^{\lambda}}_{(\alpha,f)})$ and $GVD(X^{\bar{F}^{\lambda}}_{(\alpha,f)})$ are the*

f-genetic coefficient and the f-genetic variation degree of λ order f-heredity set of $X^{\bar{F}^{\lambda}}_{(\alpha,f)}$ with regard to set $X^{\bar{F}}$ respectively, then

$$GEC(X^{\bar{F}^{\lambda}}_{(\alpha,f)}) + GVD(X^{\bar{F}^{\lambda}}_{(\alpha,f)}) = 1. \tag{14}$$

The intuitive meaning of Theorem 3.3: with the attribute complement to the attribute set α of set $X^{\bar{F}}$, the f-genetic granularity and the f-genetic coefficient decrease, while the f-genetic variation degree of f-heredity set increases. If $F = \emptyset$, then $GEC(X^{\bar{F}^{\lambda}}_{(\alpha,f)}) = 1$, and $GVD(X^{\bar{F}^{\lambda}}_{(\alpha,f)} = 0$. $GVD(X^{\bar{F}^{\lambda}}_{(\alpha,f)}) = 0$ denotes that the knowledge $X^{\bar{F}}$ has the minimum f-genetic variation at the beginning of f-heredity, which fact infers the propagate that during the propagating generation after generation, the characteristics of the interbreed species degenerate gradually.

Theorem 3.4. *(f-genetic invariability theorem) In f-heredity set, $card(X^{\bar{F}}_{\alpha})$ of f-genetic $X^{\bar{F}}_{\alpha}$ is a invariant constant, and is independent with change of genetic order λ, moreover*

$$card(X^{\bar{F}}_{\alpha}) = \eta^f, \tag{15}$$

where $\eta^f \in N^+$.

Theorem 3.4 is an intuitive fact, and the proof is omitted.

Due to definitions 3.1-3.4 and the theorems 3.1-3.4, there is the principle as follows.

The principle of f-genetic sieve $(K, G)_f$

$(K, G)_f$ is a f-genetic sieve with minimum even holes, the f-heredity set $X^{\bar{F}}_{\alpha}$ of set $X^{\bar{F}}$ could be separated from $(K, G)_f$, and other f-heredity sets are the surpluses of sieve $(K, G)_f$.

Where K is the set which composites of f-heredity set, and G is the set which composites of the granularity of f-heredity set.

4 The $\bar{f}$-Heredity of Exterior P-Set and Its Measurement

Definition 4.1. *Given general set $X = \{x_1, x_2, \cdots, x_m\} \subset U$, $\alpha = \{\alpha_1, \alpha_2, \cdots, \alpha_k\} \subset V$ is the attribute set of X, $X^F_{\alpha\backslash\{f(\alpha_i)\}}$ is called the first order $\bar{f}$-heredity set of X^F. If there is attribute $\alpha_i \in \alpha, \bar{f}(\alpha_i)\bar{\in}\alpha$, then $X^F_{\alpha\backslash\{\bar{f}(\alpha_1),\ldots,\bar{f}(\alpha_\lambda)\}}$ is called the λ order $\bar{f}$-heredity set of X^F, denoted by $X^{F^{\lambda}}_{(\alpha,\bar{f})}$, λ is called $\bar{f}$-genetic order, where $\lambda = card(\alpha^{\bar{f}})$, $\alpha^{\bar{f}} = \{f(\alpha_1), \cdots, f(\alpha_\lambda)\}$ and $\lambda \in N^+$.*

Definition 4.2. *$X_{\alpha}^{\bar{f}}$ is called the $\bar{f}$-genetic of the $\bar{f}$-heredity set $X_{(\alpha,\bar{f})_i}^{F^{\lambda}}$, if*

$$X_{\alpha}^{\bar{f}} = \bigcap_{i=1}^{t} X_{(\alpha,\bar{f})_i}^{F^{\lambda}}. \tag{16}$$

For example, $X_{(\alpha \backslash \{\bar{f}(\alpha_1), \bar{f}(\alpha_2), \bar{f}(\alpha_3)\})}^{F} = \{x_3, x_4, x_9, x_{10}, x_{17}\}$, $X_{(\alpha \backslash \{\bar{f}(\alpha_1), \bar{f}(\alpha_2)\})}^{F} = \{x_4, x_9, x_{17}\}$, $X_{(\alpha \backslash \{\bar{f}(\alpha_1)\})}^{F} = \{x_4, x_{17}\}$, so $\bar{f}$-genetic is $X_{\alpha}^{\bar{f}} = \{x_4, x_{17}\}$. The genetic orders of $\bar{f}$- heredity sets $X_{(\alpha \backslash \{\bar{f}(\alpha_1), \bar{f}(\alpha_2), \bar{f}(\alpha_3)\})}^{F}$, $X_{(\alpha \backslash \{\bar{f}(\alpha_1), \bar{f}(\alpha_2)\})}^{F}$ and $X_{(\alpha \backslash \{\bar{f}(\alpha_1)\})}^{F}$ are $\lambda = card(\{\bar{f}(\alpha_1), \bar{f}(\alpha_2), \bar{f}(\alpha_3)\}) = 3$, $\lambda = card(\{\bar{f}(\alpha_1), \bar{f}(\alpha_2)\}) = 2$, and $\lambda = card(\{\bar{f}(\alpha_1)\}) = 1$ respectively.

Definition 4.3. *$GEC(X_{(\alpha,\bar{f})}^{F^{\lambda}})$ is called $\bar{f}$-genetic coefficient of λ order $\bar{f}$-heredity set $X_{(\alpha,\bar{f})}^{F^{\lambda}}$ with regard to set X^F, moreover*

$$GEC(X_{(\alpha,\bar{f})}^{F^{\lambda}}) = GRD(X^F) \Big/ GRD(X_{(\alpha,\bar{f})}^{F^{\lambda}}). \tag{17}$$

Definition 4.4. *$GVD(X_{(\alpha,\bar{f})}^{F^{\lambda}})$ is called the $\bar{f}$-genetic variation degree of λ order $\bar{f}$-heredity set $X_{(\alpha,\bar{f})}^{F^{\lambda}}$ with regard to set X^F, moreover*

$$GVD(X_{(\alpha,\bar{f})}^{F^{\lambda}}) = 1 - GEC(X_{(\alpha,\bar{f})}^{F^{\lambda}}). \tag{18}$$

Due to the definition 4.1-4.4, the following propositions can be obtained.

Proposition 4.1. *$\bar{f}$-genetic $X_{\alpha}^{\bar{f}}$ exists in all $\bar{f}$-heredity set.*

Proposition 4.2. *Any two λ order $\bar{f}$-heredity set of X^F fulfil $X_{(\alpha,\bar{f})_i}^{F^{\lambda}} \cap X_{(\alpha,\bar{f})_j}^{F^{\lambda}} \neq \emptyset$.*

Propositions 4.1-4.2 are intuitive facts, and the proofs are omitted.

Theorem 4.1. *(The theorem of $\bar{f}$-heredity set chain) Suppose $X_{(\alpha,\bar{f})}^{F^{\lambda}}$ is the λ order $\bar{f}$-heredity set of X^F, and $\lambda = \lambda_1, \lambda_2, \cdots, \lambda_t$. If $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_t$, then*

$$X_{(\alpha,\bar{f})}^{F^{\lambda_1}} \subseteq X_{(\alpha,\bar{f})}^{F^{\lambda_2}} \subseteq \cdots \subseteq X_{(\alpha,\bar{f})}^{F^{\lambda_{t-1}}} \subseteq X_{(\alpha,\bar{f})}^{F^{\lambda_t}}. \tag{19}$$

The proof is similar to theorem 3.1, here is omitted.

Theorem 4.2. *(The maximum granularity theorem of $\bar{f}$-genetic knowledge) Suppose $X_{(\alpha,\bar{f})}^{F^{\lambda}}$ is λ order $\bar{f}$-heredity set of set X^F, If the attribute set $(\alpha \backslash \alpha^{\bar{f}})$ of $X_{(\alpha,\bar{f})}^{F^{\lambda}}$ fulfils*

$$card(\alpha \backslash \alpha^{\bar{f}}) = 1. \tag{20}$$

Then $X_{(\alpha,\bar{f})}^{F^{\lambda}}$ posses the maximum granularity, namely

$$GRD(X_{(\alpha,\bar{f})}^{F^{\lambda}}) = \max. \tag{21}$$

Proof. Because $card(\alpha\backslash\alpha^{\bar{f}}) = 1$, λ order $\bar{f}$-heredity set $X^{F^{\lambda}}_{(\alpha,\bar{f})}$ posses the maximum genetic order $\lambda_t = \max_{j=1}^{t}(\lambda_j)$. Moreover $\lambda_1 \le \lambda_2 \cdots \le \lambda_t$, so there is $card(X^{F^{\lambda_1}}_{(\alpha,\bar{f})})/card(U) \le card(X^{F^{\lambda_2}}_{(\alpha,\bar{f})})/card(U) \le \cdots \le card(X^{F^{\lambda_t}}_{(\alpha,\bar{f})})/card(U)$, It is easy to be obtained that $GRD(X^{F^{\lambda_t}}_{(\alpha,\bar{f})}) = card(X^{F^{\lambda_t}}_{(\alpha,\bar{f})})/card(U) = \max$.

Theorem 4.3. *(The relation theorem between $\bar{f}$-genetic coefficient and $\bar{f}$-genetic variation degree) Suppose $GEC(X^{F^{\lambda}}_{(\alpha,\bar{f})})$ and $GVD(X^{F^{\lambda}}_{(\alpha,\bar{f})})$ are the $\bar{f}$-genetic coefficient and the $\bar{f}$-genetic variation degree of λ order $\bar{f}$-heredity set $X^{F^{\lambda}}_{(\alpha,\bar{f})}$ with regard to knowledge X^F respectively, then*

$$GEC(X^{F^{\lambda}}_{(\alpha,\bar{f})}) + GVD(X^{F^{\lambda}}_{(\alpha,\bar{f})}) = 1. \tag{22}$$

This theorem is an intuitive fact, and the proof is omitted.

The intuitive meaning of Theorem 4.3: with the attribute α_i being removed from attribute set $\alpha = \{\alpha_1, \alpha_2, \cdots, \alpha_m\}$ one by one, $i = 1, 2, \cdots, t$; $t < m$, the $\bar{f}$-genetic coefficient $GEC(X^{F^{\lambda}}_{(\alpha,\bar{f})})$ of λ order $\bar{f}$-heredity set $X^{F^{\lambda}}_{(\alpha,\bar{f})}$ decreases gradually, while the $\bar{f}$-genetic variation degree $GVD(X^{F^{\lambda}}_{(\alpha,\bar{f})})$ increases gradually; when $card(\alpha\backslash\alpha^{\bar{f}}) = 1$, the $\bar{f}$-genetic variation degree $GVD(X^{F^{\lambda}}_{(\alpha,\bar{f})})$ attains maximum. This indicates that during the hybridization species in the reproduction, there is a degenerate phenomenon.

Theorem 4.4. *(invariance theorem of $\bar{f}$-genetic) In $\bar{f}$-heredity set, $card(X^{\overline{f}}_{\alpha})$ of $\bar{f}$-genetic $X^{F^{\lambda}}_{(\alpha,\bar{f})}$ is a constant, and is independent of the change of order λ, moreover*

$$card(X^{\bar{f}}_{\alpha}) = \eta^{\bar{f}}, \tag{23}$$

where $\eta^{\overline{f}} \in N^+$.

Due to definitions 4.1-4.4, and theorems 4.3-4.4, the principle can be obtained as follows.

The principle of $\bar{f}$-genetic sieve $(K, G)_{\bar{f}}$

$(K, G)_{\bar{f}}$ is an $\bar{f}$-genetic sieve with minimum even holes, the $\bar{f}$-genetic $X^{F^{\lambda}}_{(\alpha,\bar{f})}$ of $\bar{f}$-heredity set could be separated by $(K, G)_{\overline{f}}$, and other $\bar{f}$- heredity sets are the surpluses of sieve $(K, G)_{\overline{f}}$.

5 Conclusion

P-set was proposed in Refs.[1,2], and discussions about the characteristics of information system was presented in Refs.[3-8]. By employing the concepts in Refs.[1-8], the concept of f-heredity set of interior P-set and its measurement method are presented in this article, and the f-heredity theorems of interior P-set are proposed in this article; similarly the concept of $\bar{f}$-heredity set of exterior P-set and its measurement method are presented, and the $\bar{f}$-heredity theorems of interior P-set are proposed. It reveals the "information inertia" in system. Any new information in system depends on the old ones and its some evolution without exception, and the evolution of old information implicates heredity; the generation of the new information in system also depends on the existent of old information. Some specie in biological world, either its reproduction or sustenance of life abides by the "information inertia"law; the independence characteristic of species comes from the its "genetic code" self. A application background hides in P-set as follows: perhaps P-set and its $(f, \bar{f})$-heredity offer to people one of the helpful mathematical tool for the researches on biological heredity, species reproduction and so on.

Acknowledgements. This work is supported by the Open Project of Brain-like Key Laboratory Research of Fujian Province (BLISSOS2010105), the Natural Science Foundation of Fujian Province of China (No. 2009J01293) and the A-kind Scientific Research Item of Education Department of Fujian Province of China (No. JA09151, JA09230).

References

1. Shi, K.: P-sets and its applications. An International Journal Advances in Systems Science and Applications 2, 168–178 (2009)
2. Shi, K.: P-set. Journal of Shandong University (Natural Science) 11, 77–84 (2008)
3. Zhang, L., Ren, X.: An F-modal generated by function one direction S-rough set. Journal of Shandong University (Natural Science) 8, 77–81,89 (2008)
4. Zhang, L., Qiu, Y., et al.: Rough law generation and its separation. Journal of Shandong University (Natural Science) 3, 58–63 (2008)
5. Zhang, L., Shi, K.: Security transmission and recognition of F-knowledge. Journal of Systems Engineering and Electronics 4, 877–882 (2009)
6. Shi, K.: S-rough sets and its applications in diagnosi$S$-recognition for disease. In: IEEE Proceedings of the First International Conference on Machine Learning and Cybernetics, vol. 1, pp. 50–54 (2002)
7. Zhang, L., Qiu, Y., Ren, X.: One direction S-rough set$S$-based knowledge addition and knowledge battlement recognition. Journal of Shandong University (Natural Science) 10, 12–17 (2008)
8. Zhang, L., Cai, C., Qiu, Y.: Heredity-control based on S-rough sets. Computer Engineering and Applications 15, 216–218 (2009)

Notes on Two-Sided Quantales

Xi-juan Wang[1,2] and Wei He[3]

[1] Department of Mathematics, Nanjing University, Nanjing 210093, China
e-mail: wangxijuan69@163.com

[2] Lianyungang Teacher's College, Lianyungang 222006, China
e-mail: wangxijuan69@163.com

[3] School of Mathematics and CS., Nanjing Normal University, Nanjing 210097, China
e-mail: weihe@njnu.edu.cn

Abstract. In this paper, we show that a two-sided quantale is coherent iff it is isomorphic to the quantale of ideals of a distributive two-sided $*$-semilattice. Thus quantale representations of distributive two-sided $*$-semilattices are obtained. We also show that the free two-sided quantale generated by a set exists. And then the category **TQuant** of two-sided quantales is an algebraic category.

Keywords: Two-sided quantales, $*-$semilattice, ideal, homomorphism.

1 Introduction and Preliminaries

The concept of quantale was introduced by C.J.Mulvey in [7] with the purpose of studying the spectrum of c*-algebra, as well as constructive foundations for quantum mechanics. There are abundant contents in the structure of quantales, because quantales can be regarded as the generalization of the notion of complete Heyting algebra(cHa). The research of quantales has related to several research areas such as non-commutative c*-algebra, the ideal theory of rings, linear logic, theoretic computer science and the sheaf theory (see [1], [2],[5]). In this paper we investigate a special class of quantales—-two-sided quantales. We show that a two-sided quantale is coherent iff it is isomorphic to the quantale of ideals of a distributive two-sided $*$-semilattice, thus, a quantale representation of distributive two-sided $*$-semilattice is obtained. We also show that the free two-sided quantale generated by a set exists, and then, the category **TQuant** is an algebraic category.

A quantale is a complete lattice Q together with an associative binary operation $*$ satisfying

$$a*(\bigvee b_s) = \bigvee(a*b_s) \text{ and } (\bigvee b_s)*a = \bigvee(b_s*a) \text{ for all } a \in Q \text{ and } \{b_s\} \subseteq Q.$$

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 745–753.
springerlink.com

Let Q be a quantale. Q is commutative if $*$ is commutative. An element a of Q is called left-sided if $1 * a = a$, where 1 is the largest element of Q. If every element of Q is left-sided, then Q is called a left-sided quantale. The notions of right-sided quantales and two-sided quantales are defined similarly.

A typical example of commutative two-sided quantales is the poset Idl(R) of two-sided ideals of a commutative ring R by defining $I_1 \& I_2 = \{a_1a_2 | a_1 \in I_1, a_2 \in I_2\}$.

Definition 1.1. Let L be a sup-semilattice with the largest element 1, and $*$ be an associative binary operation on L. Then L is called a $*$-semilattice.

Let L be a $*$-semilattice. We call a a prime element of L if it satisfies: $a \neq 1$ and $x * y \leq a$ implies $x \leq a$ or $y \leq a$ for any $x, y \in L$. If L satisfies: $a*(b_1 \vee b_2) = (a*b_1) \vee (a*b_2)$, $(b_1 \vee b_2)*a = (b_1*a) \vee (b_2*a)$ and $a*0 = 0*a = 0$ for any $a, b_1, b_2 \in L$, then we call L a distributive $*$-semilattice. If L satisfies: $a * 1 = a = 1 * a$ for all $a \in L$, then we call L a two-sided $*$-semilattice. Obviously a quantale is a $*$-semilattice, and also a distributive $*$-semilattice. A two-sided quantale is a two-sided $*$-semilattice.

Let Q and Q' be quantales. A quantale homomorphism $f : Q \to Q'$ is a $\vee$,$*$ and 1 preserving function. The category of quantales and quantale homomorphisms is denoted by **Quant** and the category of two-sided quantales denoted by **TQuant**. Let L_1 and L_2 be $*$-semilattices. A $*$-semilattice homomorphism $f : L_1 \to L_2$ is a $\vee$,$*$ and 1 preserving function. The category of distributive two-sided $*$-semilattices and $*$-semilattice homomorphisms is denoted by $*$**DTSLAT**.

Definition 1.2. *Let L be a $*$-semilattice and $\emptyset \neq I \subseteq L$. We call I a right-ideal if I satisfies the following conditions:*

(1) *If $a \in I$ and $b \in I$, then $a \vee b \in I$;*

(2) *If $a \in I$ and $r \in L$, then $a * r \in I$;*

(3) *If $a \in I$ and $b \leq I$, then $b \in I$.*

Similarly we can define the concept of left-ideal. If I is a right-ideal and also is a left-ideal, then we call it an ideal. In a distributive two-sided $*$-semilattice, $\downarrow a$ is an ideal. The set of all ideals of L is denoted by $Idl(L)$.

Definition 1.3. *Let L be a $*$-semilattice and $\emptyset \neq F \subseteq L$. We call F a filter if F satisfies:*

(1) *If $a \in F$ and $b \in F$, then $a * b \in F$;*

(2) *F is an upper set.*

The set of all filters of L is denoted by $Fil(L)$.

Definition 1.4. *Let L be a $*$-semilattice and I be a proper ideal of L. We call I a prime ideal if it satisfies: $a * b \in I$ implies $a \in I$ or $b \in I$. Dually a proper filter F of L is called to be a prime filter if it satisfies: $a \vee b \in F$ implies $a \in F$ or $b \in F$.*

2 The Quantale $\boldsymbol{Idl(L)}$

We regard the two-elements lattice $\{0,1\}$ as a $*$-semilattice by taking $* = \wedge$.

Lemma 2.1. *Let L be a $*$-semilattice and $f : L \to \{0,1\}$ be a $*$-semilattice homomorphism. Then $f^{-1}(0) = \{a \in L | f(a) = 0\}$ is an ideal of L and $f^{-1}(1) = \{a \in L | f(a) = 1\}$ is a filter.*

Proof. Obviously $f^{-1}(0) \neq \emptyset$. For any $a, b \in f^{-1}(0)$, $f(a) = 0$ and f(b)=0. Since f is a $*$-semilattice homomorphism, $f(a \vee b) = f(a) \vee f(b) = 0 \vee 0 = 0$, and hence $a \vee b \in f^{-1}(0)$. For any $r \in L$ and $a \in f^{-1}(0)$, $f(a * r) = f(a) * f(r) = 0 * f(r) = 0$ and $f(r * a) = f(r) * f(a) = f(r) * 0 = 0$ since f is a homomorphism. So $a * r \in f^{-1}(0)$ and $r * a \in f^{-1}(0)$. f preserves order since it is a homomorphism, then if $f(a) = 0$ and $b \leq a$, then $f(b) = 0$.

Obviously $f^{-1}(1) \neq \emptyset$. Since f preserves finite sups, f preserves order, and hence $f^{-1}(1)$ is an upper set. For any $b_1, b_2 \in f^{-1}(1)$, $f(b_1) = f(b_2) = 1$. Then $f(b_1 * b_2) = f(b_1) * f(b_2) = 1 * 1 = 1$. So $b_1 * b_2 \in f^{-1}(1)$. Thus $f^{-1}(1)$ is a filter.

Proposition 2.2. *Let L be a $*$-semilattice and $I \subseteq L$ be an ideal. Then the following statements are equivalent:*

(1) *$L \setminus I$ is a (prime) filter of L.*

(2) *I is a prime ideal of L.*

(3) *There exists a $*$-semilattice homomorphism $f : L \to \{0,1\}$ such that $f^{-1}(0) = I$.*

(4) *There exists a $*$-semilattice homomorphism $f : L \to \{0,1\}$ such that $f^{-1}(1) = L \setminus I$.*

Proof. (1) Implies (2): Suppose $F = L - I$ be a filter of L. Then $1 \in F$, and hence $1 \notin I$. For any $a, b \in L$ with $a * b \in I$, $a * b \notin F$. If $a \notin I$ and $b \notin I$, then $a, b \in F$. Since F is a filter, $a * b \in F$. A contradiction to $a * b \notin F$. So $a \in I$ or $b \in I$. Then I is a prime ideal.

(2) Implies (3): Define $f : L \to \{0,1\}$ as the following: for any $a \in L$, $f(a) = 0$ if $a \in I$ and $f(a) = 1$ if $a \notin I$. Then $f^{-1}(0) = I$. Since I is prime, $I \neq L$ and $I \neq \emptyset$. So $f(1) = 1$ and $f(0) = 0$. For any $a_1, a_2 \in L$, $f(a_1 \vee a_2) = f(a_1) \vee f(a_2)$ can be directly checked. If $a_1 * a_2 \in I$, then $a_1 \in I$ or $a_2 \in I$ since I is a prime ideal. So $0 = f(a_1 * a_2) = f(a_1) * f(a_2)$. If $a_1 * a_2 \notin I$, then $a_1 \notin I$ and $a_2 \notin I$ since I is an ideal. So $1 = f(a_1 * a_2) = f(a_1) * f(a_2)$.

(3) Implies (1): Suppose $f : L \to \{0,1\}$ be a $*$-semilattice homomorphism and $f^{-1}(0) = I$. Then f is surjective, and hence $f^{-1}(1) = L - I$. By Lemma 2.1, $f^{-1}(1) = L - I$ is a filter. Since $0 \notin L - I$, $L - I$ is a proper filter of L. For any $a, b \in L$ with $a \vee b \in L - I$, $f(a \vee b) = 1$. Since f is a $*$-semilattice homomorphism, $f(a) \vee f(b) = 1$. Then $f(a) = 1$ or $f(b) = 1$, i.e. $a \in L - I$ or $b \in L - I$. Thus $L - I$ is a prime filter.

The equivalence of (3) and (4) is obvious.

Theorem 2.3. *Let L be a distributive two-sided $*$-semilattice and $I \in Idl(L), F \in Fil(L)$. If I is maximal for that $I \cap F = \emptyset$, then I is a prime ideal.*

Proof. Since $1 \in F$, I is a proper ideal. Suppose $a_1 * a_2 \in I$. Denote the ideals generated by I and a_i(i=1,2) by K_i. Then $K_i =\downarrow \{x \vee b | x \in I, b \in L, b \leq a_i\}$, i=1,2. If $K_i \cap F \neq \emptyset$, i=1,2, then there exist $x_1, x_2 \in I$ and $b_1 \leq a_1, b_2 \leq a_2$ such that $x_1 \vee b_1 \in F$ and $x_2 \vee b_2 \in F$. Since F is a filter, $d = (x_1 \vee b_1) * (x_2 \vee b_2) = (x_1 * x_2) \vee (x_1 * b_2) \vee (b_1 * x_2) \vee (b_1 * b_2) \in F$. But $x_1 * x_2$, $x_1 * b_2$, $b_1 * x_2$ and $b_1 * b_2$ all belong to I since I is an ideal and $a_1 * a_2 \in I$, and hence $d \in I$. Then $I \cap F \neq \emptyset$, a contradiction. So either $K_1 \cap F = \emptyset$ or $K_2 \cap F = \emptyset$, i.e. $a_1 \in I$ or $a_2 \in I$. So I is a prime ideal.

Let L be a distributive $*$-semilattice. Then $Idl(L)$ is a complete lattice under inclusion order. For a family of ideals $\{I_\lambda : \lambda \in \Lambda\}$ of L, $\bigvee I_\lambda =\downarrow \{x_1 \vee x_2 \vee \cdots \vee x_n | x_i \in I_{\lambda_i}, \lambda_i \in \Lambda, i = 1, 2 \cdots, n.\}$. For $I, J \in Idl(L)$, define $I \& J =\downarrow \{a_i * a_j | a_i \in I, a_j \in J\}$. We easily check that $I \& J$ is an ideal of L, and then $Idl(L)$ is a quantale. If L is a distributive two-sided $*$-semilattice, then $Idl(L)$ is a two-sided quantale. Specially, if L is a two-sided quantale, then $Idl(L)$ is a two-sided quantale.

Proposition 2.4. *Let L be a distributive two-sided $*$-semilattice. Then a prime element of $Idl(L)$ just is a prime ideal of L.*

Proof. Suppose I be a prime element of $Idl(L)$ and $a * b \in I$. obviously $1 \notin I$. Write $J =\downarrow a$, $K =\downarrow b$, then $J \& K =\downarrow (a * b)$, and hence $J \& K \subseteq I$. Since I is a prime element of $Idl(L)$, $\downarrow a \subseteq I$ or $\downarrow b \subseteq I$. Then $a \in I$ or $b \in I$. Conversely, suppose I be a prime ideal of L. Then $I \neq L$. Let $J, K \in Idl(L)$ and $J \& K \subseteq I$. If $K \subseteq I$ is not true, then there exists $a \in K \setminus I$. For any $b \in J$, $a * b \in J \& K \subseteq I$, i.e. $a * b \in I$. Since $a \notin I$ and I is a prime ideal, $b \in I$. Then $J \subseteq I$. Similarly, If $J \subseteq I$ is not true, then $K \subseteq I$. So I is a prime element of $Idl(L)$.

An element a of a quantale L is called a finite element if it satisfies the condition: for every $S \subseteq L$ with $a \leq \bigvee S$, there exists a finite $F \subseteq S$ with $a \leq \bigvee F$. The set of all finite elements of L is denoted by $K(L)$. We know $K(L)$ is a sub-$\vee$- semilattice of L. We define a quantale L to be coherent if it satisfies: (i) every element of L can be expressed as a join of finite elements, and (ii) the finite elements form a sub $*$-semilattice of L—equivalently, $a * b$ is finite for any finite elements a, b and 1 is finite.

Lemma 2.5. *Let L be a distributive two-sided $*$-semilattice. Then $I \in Idl(L)$ is finite if and only if it is a principle ideal of L, i.e. $I =\downarrow a$ for some $a \in L$.*

Proof. Suppose $I \in Idl(L)$ is finite. Then $I = \bigcup\{\downarrow x | x \in I\} = \bigvee_{Idl(L)}\{\downarrow x | x \in I\}$. Since I is finite and $\downarrow x$ is an ideal, there exist $x_1, x_2, \cdots, x_n \in I$ such that $I \subseteq \bigvee_{Idl(L)}\{\downarrow x_i | i = 1, 2, \cdots, n.\} =\downarrow (\vee_{i=1}^n x_i)$. $\downarrow (\vee_{i=1}^n x_i) \subseteq I$ is obviously, and then $\downarrow (\vee_{i=1}^n x_i) = I$, this shows I is a principle ideal.

Conversely, suppose I be a principle ideal of L, i.e. there is $a \in L$ such that $I = \downarrow a$. Let $\{I_\lambda : \lambda \in \Lambda\} \subseteq Idl(L)$ with $\downarrow a \subseteq \bigvee_{Idl(L)} I_\lambda$. Then $a \in \bigvee_{Idl(L)} I_\lambda$, and hence there exist a natural number n and $x_i \in I_{\lambda_i}, i = 1, 2, \cdots, n$ such that $a \leq x_1 \vee x_2 \vee \cdots \vee x_n$. Then $a \in \bigvee_{i=1}^{i=n} I_{\lambda_i}$, and then $\downarrow a \subseteq \bigvee_{i=1}^{i=n} I_{\lambda_i}$, this is to say that $I = \downarrow a$ is a finite element of $Idl(L)$.

Theorem 2.6. *A two-sided quantale is coherent iff it is isomorphic to the quantale of ideals of a distributive two-sided $*$-semilattice.*

Proof. Suppose L be a two-sided quantale and it is isomorphic to the quantale of ideals of a distributive two-sided $*$-semilattice D. By Lemma 2.5, $K(Idl(D)) = \{\downarrow a | a \in D\}$. For any $a, b \in D$, $(\downarrow a)\&(\downarrow b) = \downarrow (a * b)$ and for any $I \in Idl(D)$, $I = \bigvee_{Idl(D)}\{\downarrow x | x \in I\}$. So $Idl(L)$ is coherent and then L is coherent. Conversely, suppose L be a two-sided and coherent quantale. Then $K(L)$ is a distributive two-sided $*$-semilattice. Define $f : L \to Idl(K(L))$ by $f(a) = \{k \in K(L) | k \leq a\}$ for any $a \in L$. We easily check that the definition of f is suitable. For any $I \in Idl(K(L))$, it is a directed subset of L, and so $k \in K(L)$, $k \leq \bigvee_L I$ implies $k \leq i$ for some $i \in I$, and hence $k \in I$. Thus $f(\bigvee_L I) = I$ for all $I \in Idl(K(L))$; but the condition (i) in the definition of coherence tells us that $\bigvee_L (f(a)) = a$ for all $a \in L$. So f is a bijection(clearly order preserving) between L and $Idl(K(L))$. Then f preserves arbitrary sups. It remains to check that $f(a * b) = f(a)\&f(b)$ for any $a, b \in L$. $f(a * b) \supseteq f(a)\&f(b)$ is obvious. For any $k \in K(L)$ with $k \leq a * b$, $k \leq (\bigvee_L\{k_a \in K(L) | k_a \leq a\}) * (\bigvee_L\{k_b \in K(L) | k_b \leq b\}) = \bigvee_L\{k_a * k_b | k_a, k_b \in K(L) \text{ and } k_a \leq a, k_b \leq b\}$. Since $k \in K(L)$, there exist a natural number n and $k_{a_i}, k_{b_i} \in K(L)$ with $k_{a_i} \leq a$ and $k_{b_i} \leq b, i = 1, 2, \cdots, n$ such that $k \leq \vee_{i=1}^{n}(k_{a_i} * k_{b_i})$. Then $k \in f(a)\&f(b)$, and hence $f(a * b) \subseteq f(a)\&f(b)$.

3 Free Two-Sided Quantales

Definition 3.1. *Let $(A, \leq)$ be a poset with the largest element 1, and $*$ be an associative binary operation on L. We call L a $*$-algebra if it satisfies: $a_1 \leq a_2$ implies $a * a_1 \leq a * a_2$ and $a_1 * a \leq a_2 * a$ for any $a, a_1, a_2 \in A$.*

We call a $*$-algebra A a two-sided $*$-algebra if it satisfies: $a * 1 = a = 1 * a$ for all $a \in A$. And we call a two-sided $*$-algebra A a commutative two-sided $*$-algebra if $*$ is commutative.

Obviously a quantale is a $*$-algebra, a two-sided quantale is a two-sided $*$-algebra. But the converse is not true. The interval $(-\infty, 1]$ is not a complete lattice, and then it is not a quantale. But it is a two-sided $*$-algebra by defining $* = \wedge$.

Let A and A' be $*$-algebras. A $*$-algebra homomorphism $f : A \to A'$ is a $*$,1 and order preserving function. The category of two-sided $*$-algebras and $*$-algebra homomorphisms is denoted by $*$**TAlg** and the category of commutative two-sided $*$-algebras and $*$-algebra homomorphisms is denoted by $*$**CTAlg**.

Let A be a set. The set of all finite sequences $\{(a_1, a_2, \cdots, a_n)|a_i \in A, i = 1, 2, \cdots, n.n = 1, 2, \cdots\}$ (including empty sequence) is denoted as ωA. We define an order $\leq$ on ωA by $\omega_1 = (a_1, a_2, \cdots, a_n) \leq \omega_2 = (b_1, b_2, \cdots, b_m)$ iff ω_2 is a subsequence of ω_1, and define an associative binary operation $*$ on ωA by $\omega_1 * \omega_2 = (a_1, a_2, \cdots, a_n, b_1, b_2, \cdots, b_m)$ for any $\omega_1 = (a_1, a_2, \cdots, a_n), \omega_2 = (b_1, b_2, \cdots, b_m)$. Then ωA is a two-sided $*$-algebra.

Lemma 3.2. *The free two-sided $*$-algebra generated by a set A is the $*$-algebra $(\omega A, \leq, *)$.*

Proof. The unit map $\eta_A : A \to \omega A$ sends $a \in A$ to the singleton sequence (a). Now any $\omega = (a_1, a_2, \cdots, a_n) \in \omega A$ can be uniquely expressed as $(a_1) * (a_2) * \cdots * (a_n)$. So any map $f : A \to B$, where B is a two-sided $*$-algebra, can uniquely extended to a $*$-algebra homomorphism

$\bar{f} : \omega A \to B$; $\omega = (a_1, a_2, \cdots, a_n) \mapsto f(a_1) * f(a_2) * \cdots * f(a_n)$.

For $S \in ob(*\mathbf{TAlg})$, take $\mathfrak{D}S = \{A \subseteq S | A =\downarrow A\}$ and ordered by inclusion. Then $\mathfrak{D}S$ is a complete lattice. For any $A_1, A_2 \in \mathfrak{D}S$, define $A_1 \& A_2 =\downarrow \{a_1 * a_2 | \, a_1 \in A_1, a_2 \in A_2\}$ and $A_1 \& \emptyset = \emptyset \& A_1 = \emptyset$. We can easily verify that $\mathfrak{D}S$ is a two-sided quantale.

For a $*$-algebra homomorphism $h : S \to T$ define $\mathfrak{D}h = (A \mapsto \downarrow h(A))$: $\mathfrak{D}S \to \mathfrak{D}T$. Obviously $\mathfrak{D}h$ preserves arbitrary $\vee$ and 1. For any $A_1, A_2 \in \mathfrak{D}S$, $\mathfrak{D}h(A_1 \& A_2) =\downarrow h(A_1 \& A_2) =\downarrow h(\downarrow \{a_1 * a_2 | \, a_1 \in A_1, a_2 \in A_2\})$; $\mathfrak{D}h(A_1) \& \mathfrak{D}h(A_1) = [\downarrow h(A_1] \& [\downarrow h(A_2)] =\downarrow \{b_1 * b_2 | \, b_1 \leq h(a_1), b_2 \leq h(a_2), a_1 \in A_1, a_2 \in A_2\}$. Let $b \leq b_1 * b_2$, where $b_1 \leq h(a_1), b_2 \leq h(a_2), a_1 \in A_1, a_2 \in A_2\}$. Then $b \leq h(a_1) * h(a_2) = h(a_1 * a_2)$, and hence $b \in \mathfrak{D}h(A_1 \& A_2)$. On the other hand, let $d \leq h(b)$, where $b \leq a_1 * a_2$. Then $h(b) \leq h(a_1 * a_2) = h(a_1) * h(a_2)$ since h is a $*$-algebra homomorphism. So $d \in \mathfrak{D}h(A_1) \& \mathfrak{D}h(A_1)$. Thus $\mathfrak{D}h(A_1 \& A_2) = \mathfrak{D}h(A_1) \& \mathfrak{D}h(A_1)$. The above shows $\mathfrak{D}h$ is a quantale homomorphism.

Lemma 3.3. $\mathfrak{D}$*: $*$**TAlg** $\to$ **TQuant** is a functor.*

Theorem 3.4. *The functor $\mathfrak{D}$: $*$**TAlg** $\to$ **TQuant** is a left adjoint to the forgetful functor $\mathfrak{U}$: **TQuant** $\to *$**TAlg**.*

Proof. The unit map $\eta_S : S \to \mathfrak{U}\mathfrak{D}S$ sends $a \in S$ to the down set $\downarrow a$. Easily see that η_S is a $*$-algebra homomorphism.

For any $*$-algebra homomorphism $f : S \to B$, where B is a two-sided quantale, it has unique extension to a quantale homomorphism: $\bar{f} : \mathfrak{U}\mathfrak{D}S \to B$; $A \mapsto \bigvee\{f(a) | \, a \in A\}$. Obviously, $f = \bar{f} \circ \eta_S$ and $\bar{f}$ preserves arbitrary $\vee$ and 1. It remains to show that $\bar{f}$ preserves &. For any $A_1, A_2 \in \mathfrak{D}S$, $\bar{f}(A_1) * \bar{f}(A_2) = (\bigvee\{f(a_1) | \, a_1 \in A_1\}) * (\bigvee\{f(a_2) | \, a_2 \in A_2\}) = \bigvee\{f(a_1) * f(a_2) | \, a_1 \in A_1, \, a_2 \in A_2\} = \bigvee\{f(a_1 * a_2) | \, a_1 \in A_1, \, a_2 \in A_2\}$; $\bar{f}(A_1 \& A_2) = \bar{f}(\downarrow \{a_1 * a_2 | \, a_1 \in A_1, \, a_2 \in A_2\}) = \bigvee\{f(b) | \, b \leq a_1 * a_2, \, a_1 \in A_1, \, a_2 \in A_2\}$. Then $\bar{f}(A_1) * \bar{f}(A_2) = \bar{f}(A_1 \& A_2)$.

By Lemma 3.2 and Theorem 3.4 ,we have

Theorem 3.5. *The free two-sided quantale generated by a set A is* $(\mathfrak{D}(\omega A)$, $\subseteq$, &$)$.

Corollary 3.6. *The category **TQuant** is an algebraic category.*

4 The Coproduct of Commutative Two-Sided Quantales

In this section we discuss the coproduct of commutative two-sided quantales, and we write **CTQuant** for the category of commutative two-sided quantales and quantale homomorphisms.

Let S be a $*$-algebra. By a coverage on S we mean a function C assigning to each $s \in S$ a set $C(s)$ of subsets of $\downarrow s$, called covering of s, with the $*-$stability property: $A \in C(s) \Rightarrow \{a * r | \, a \in A\} \in C(s * r)$ and $\{r * a | \, a \in A\} \in C(r * s)$ for any $r \in S$.

Given a coverage C on a $*$-algebra S, we call a subset I of S to be $C-ideal$ if it is a lower set and satisfies: $(\exists A \in C(a))(A \subseteq I) \Rightarrow a \in I$ for all $a \in S$. We write **C-Idl(S)** for the set of all $C - ideals$ of S, ordered by inclusion.

For example, if S is a distributive two-sided $*$-semilattice(as an $*$-algebra), we could take $C(s)$ to be the set of all finite sets with join s and the $C-ideals$ of S are just its *ideals*. $*-$stability in this case is just the distributive law. By a site we mean a $*$-algebra equipped with a coverage. We say a quantale B is freely generated by a site (S, C) if there is a $*$-algebra homomorphism $f : S \to B$ which transforms covers to joins in the sense that for every $s \in S$ and every $A \in C(s)$ we have $f(s) = \vee_B\{f(a) | \, a \in A\}$, and which is universal among such maps, i.e. every $f^{'} : S \to B^{'}$ satisfying the same conditions factors uniquely through f by a quantale homomorphism $B \to B^{'}$.

Let Q be a quantale. A quantic nucleus on Q is a closure operator j such that $j(a) * j(b) \leq j(a * b)$, for all $a, b \in Q$, and if j is a closure operator on a complete lattice Q, then the image $Q_j = \{a \in Q | \, j(a) = a\}$ is complete, where the join in Q_j is given by $\vee_j(a_\gamma) = j(\vee a_\gamma)$.

Lemma 4.1*(see [2]). If $j : Q \to Q$ is a quantic nucleus, then Q_j ia a quantale via $a *_j b = j(a * b)$ for any $a, b \in Q_j$; and $j : Q \to Q_j$ is a quantale homomorphism. We call Q_j a quotient quantale.*

Theorem 4.2. *Let S be a $*$-algebra. For any site (S, C), **C-Idl(S)** is a quantale, and is freely generated by (S, C).*

Proof. First we show that **C-Idl(S)** is a quotient quantale of the quantale $\mathfrak{D}S$. It is clear that an arbitrary intersection of $C-ideals$ is a $C-ideal$. If we define $j : \mathfrak{D}S \to \mathfrak{D}S$ by $j(A) = \bigcap\{I \in C - Idl(S) | \, I \supseteq A\}$ for any $A \in \mathfrak{D}S$, then we have $A \subseteq j(A) = j(j(A))$ for any $A \in \mathfrak{D}S$, and the image of j is precisely **C-Idl(S)**. So we need only to show that $j(A_1)\&j(A_2) \subseteq j(A_1\&A_2)$ for any $A_1, A_2 \in \mathfrak{D}S$.

Write I for $j(A_1\&A_2)$. Consider $J = \{s \in S|\,(\forall a_1 \in A_1))(a_1 * s \in I)\}$; It is clear that $A_2 \subseteq J$ since $A_1\&A_2 \subseteq I$. We shall show that J is a $C-ideal$. J is obviously a lower set. Suppose $U \in C(s)$, $U \subseteq J$, then for every $a_1 \in A_1$ we have $\{a_1 * u|\, u \in U\} \in C(a_1 * s)$ by $*$-stability of C, and $\{a_1 * u|\, u \in U\} \subseteq I$ by the definition of J. Since I is a $C - ideal$, we deduce $a_1 * s \in I$ for all $a_1 \in A_1$, and hence $s \in J$.

Now if we define $K = \{s \in S|\,(\forall t \in J)(s * t \in I)\}$, then a similar argument shows K is a $C - ideal$, and $A_1 \subseteq K$ since $A_1\&J \subseteq I$. Then we have $j(A_1)\&j(A_2) \subseteq K\&J \subseteq I = j(A_1\&A_2)$, So j is a nucleus and **C-Idl(S)** is a quotient quantale of $\mathfrak{D}S$.

We write f for the composite map:

$$S \xrightarrow{\downarrow(-)} \mathfrak{D}S \xrightarrow{j} (\mathfrak{D}S)_j = \textbf{C-Idl(S)}.$$

For any $A \in C(s)$, since $A \subseteq \bigcup\{\downarrow a|\, a \in A\} \subseteq j(\bigcup\{\downarrow a|\, a \in A\})$, we have $s \in j(\bigcup\{\downarrow a|\, a \in A\})$, so f transforms covers to joins. Let $g : S \to B$ be any other $*$-algebra homomorphism from S to a quantale with this property, it is easy to verify that the right adjoint $h : B \to \mathfrak{D}S$ of the unique extension $\bar{g} : DS \to B$ of g to a quantale homomorphism is given by $h(b) = \{s \in S|\, f(s) \leq b\}$, and since g transforms covers to joins this set is a $C - ideal$, i.e. h factors through **C-Idl(S)**. So $\bar{g}$ factors (uniquely) through $j : \mathfrak{D}S \to$ **C-Idl(S)**, i.e. there is an unique mapping $g^* :$ **C-Idl(S)** $\to B$ such that $\bar{g} = g^* \circ j$. Since $\bar{g}$ is a quantale homomorphism and j is surjective, g^* is a quantale homomorphism.

Let L_1 and L_2 be two-sided coherent quantales. Then by Theorem 2.6 L_i is freely generated by $K(L_i)$,i=1,2. So any $*-$semilattice homomorphism $K(L_1) \to K(L_2)$ extends uniquely to a quantale homomorphism $L_1 \to L_2$. We define that a quantale homomorphism $f : L_1 \to L_2$ between coherent quantales is coherent if f maps $K(L_1)$ into $K(L_2)$. Then we have

Corollary 4.3. *The category* $*$***DTSLAT*** *is dual to the category* ***TCohQ*** *of two-sided coherent quantales and coherent maps between them.*

We conclude this section with an application of Theorem 4.2: to obtain the construction of coproducts of **CTQuant**.

Let $(Q_\gamma|\, \gamma \in \mathit{\Gamma})$ be a family of commutative two-sided quantales, and write B for the set theoretic product of the Q_γ with a $*$ operation defined by $(a_\gamma)_{\gamma\in\mathit{\Gamma}} * (b_\gamma)_{\gamma\in\mathit{\Gamma}} = (a_\gamma * b_\gamma)_{\gamma\in\mathit{\Gamma}}$ (of course B is a commutative two-sided quantales, and it is the product of Q_λ in **CTQuant**). For each γ, the projection $p_\gamma : B \to Q_\gamma$ has a right adjoint $q_\gamma : Q_\gamma \to B$, which sends $a \in Q_\gamma$ to the unique element b with $p_\gamma(b) = a$ and $p_\delta(b) = 1$ for $\delta \neq \gamma$; and easily check that q_γ preserves $*$. Let Q be the sub-$*-$algebra of B generated by the union of the images of the q_γ, i.e., the set of all $b \in B$, such that $p_\gamma(b) = 1$ for all but a finite number of indices γ; then it is easy to see that the maps $q_\gamma : Q_\gamma \to Q$ make Q into the coproduct of the Q_γ in $*$**CTAlg**.

From the universal property of coproducts, it ie clear that we should have a $*-$algebra homomorphism from Q to the coproduct of Q_γ in $*$**CTQuant**, which is universal among homomorphisms f such that each of the composites $f \circ q_\gamma$ preserves joins. We accordingly define a coverage C on Q, as follows: if $a \in Q$ and $S \subseteq Q_\gamma$, define $S[\gamma, a]$ to be the set of all elements of Q obtained on replacing the γth entry of a by a member of S. Then define

$C(a) = \{S[\gamma, a] | \, \gamma \in \Gamma \text{ and } S \subseteq Q_\gamma \text{ and } \vee S = p_\gamma(a)\}$.

It is easily verified that C satisfies the $*$-stability condition, and that a $*$-algebra homomorphism f from Q to a quantale transforms covers in C to joins iff each of the composites $f \circ q_\gamma$ preserves joins. So we have

Theorem 4.4. *The coproduct of* $\{Q\gamma | \, \gamma \in \Gamma\}$ *in the category* ***CTQuant*** *is* $C - Idl(Q)$.

Acknowledgements. This work described here is partially supported by the grants from the National Natural Science Foundation of China (NSFC No. 10331011).

References

1. Borceux, F., Van Den Bossche, G.: An essay on non-commutative topology. Topology Appl. 31, 203–223 (1989)
2. Niefield, S.B., Rosenthal, K.I.: Constructing from quantales. Mathematical Proceedings of the Cambridge Philosophical Society 104, 215–233 (1988)
3. Johnstone, P.T.: Stone spaces. Cambridge University Press, New York (1989)
4. Wei, H.: Category theory. Science Press, Beijing (1989)
5. Borceux, F., Van Den Bossche, G.: Quantales and their sheaves. Oder. 3, 61–87 (1986)
6. Chongyou, Z., lei, F., Hongbin, C.: Frames and continuous lattices. Capital Normal University Press, Beijing (2000)
7. Mulvey, C.J., Rend: Circ.Mat.Palermo, vol. 12, pp. 99–104 (1986)

Counting the Solutions of a System of Equations over Finite Field is NP-Hard

Hong Lin

School of Sciences, Jimei University, Xiamen 361021, P. R. China
e-mail: linhongjm@163.com

Abstract. Let G be a simple connected graph with r vertices and t edges. Let d be a positive integer and $a \in F^*_{p^n}$. The associated system of equations $S(G, d, p^n, a)$ of G over a finite field F^n_p is defined to be $AX = C$, where $X = (x^d_1, x^d_2, ..., x^d_t)^\top$, $C = (a, a, ..., a)^\top$, $A = (a_{ij})$ is the incidence matrix of G. We show that the solutions of $S(G, d, p^n, a)$ has a relation with the number of perfect matchings of G. We also show that the problem of determining the number of solutions to a system of equations over finite field is NP-hard.

Keywords: NP-hard Problem, perfect matching, finite field.

1 Introduction

All graphs considered in this paper will be simple graphs. Let G be a graph with vertex set $V(G)$ and edge set $E(G)$. A *matching* M of G is a subset of $E(G)$ such that any two edges of M have no vertices in common [6]. A *matching* M is said to be a *perfect matching* if it covers all vertices of G. The number of perfect matchings of G is denoted by $\phi(G)$. From the definition of perfect matching, we know that if G has a perfect matching, then the number of vertices of G must be even. Let a, b be two positive integers, the largest common divisor of a and b is denoted by $gcd(a, b)$.

Let F^n_p denote a finite field with p^n elements, where p is a prime, and let $F^*_{p^n} = F_{p^n} \setminus \{0\}$. Finite field theory has many applications in theoretical computer science, especially in coding theory [5]. Estimating the number of solutions of equations over a finite field is a charming and important problem with a long history. As mentioned in [3], C.F.Gauss first studied the solutions of equation $ay^l + bz^m + c = 0$ over a a finite field, and calculated the number of solutions for $(l, m) = (2, 2), (3, 3), (4, 4), (2, 4)$. A. Weil [8] made a further study on solutions of the equation $a_0x^{n_0}_0 + a_1x^{n_1}_1 + \cdots + a_rx^{n_r}_r = 0$, and posed a famous conjecture which has been very influential in the recent development of both number theory and algebraic geometry. Deligine finally solved Weil's conjecture and was awarded the Fields Medal in 1978 [4]. For some more

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 755–759.
springerlink.com

results on this topic, we refer to [5, 7, 8, 9]. Other graph, finite field and algorithm theoretical terminologies and notations not defined in this paper we refer the readers to [1, 5, 2].

Let G be a graph with r vertices $v_1, v_2, \cdots, v_r$ and t edges $e_1, e_2, \cdots, e_t$. Then the *incidence matrix* A of G is defined as a $r \times t$ matrix $A=(a_{ij})$, where $a_{ij} = 1$ if v_i and e_j are incident, or $a_{ij} = 0$ if v_i and e_j are not incident. Let d be a positive integer and $a \in F^*_{p^n}$. Now we can define an *associated system of equations* $S(G, d, p^n, a)$ of G over the finite field F_{p^n} as following

$$AX = C,$$

where $X = (x_1^d, x_2^d, ..., x_t^d)^\top$, $C = (a, a, ..., a)^\top$, $A = (a_{ij})$ is the incidence matrix of G.

For an example, let G be a cycle with 4 vertices v_1, v_2, v_3, v_4 and 4 edges e_1, e_2, e_3, e_4. (see Figure 1)

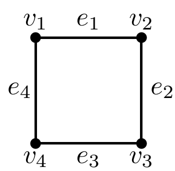


Fig. 1.

Set $d = 2$, $p = 3$, $n = 1$, $a = 1$. Then $X = (x_1^2, x_2^2, x_3^2, x_4^2)^\top$, the incidence matrix A of the cycle G is

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix},$$

and the associated system of equations $S(G, 2, 3, 1)$ over the finite field F_3 of the cycle G is

$$\begin{cases} x_1^2 + x_4^2 = 1, \\ x_1^2 + x_2^2 = 1, \\ x_2^2 + x_3^2 = 1, \\ x_3^2 + x_4^2 = 1. \end{cases}$$

The number of solutions of an associated system of equations $S(G, d, p^n, a)$ of a graph G is denoted by $N(S(G, d, p^n, a))$.

In this paper, we show that the solutions of an associated system of equations $S(G, d, p^n, a)$ has a relation with the number of perfect matching of G. More precisely, we can give some lower bounds of $N(S(G, d, p^n, a))$ by $\phi(G)$. Moreover, we also prove that the problem of determining the number of solutions to a system of equations over finite field is $NP - hard$.

2 Main Results

To prove our main results, we need the following lemma.

Lemma 1 ([3]). *Let $a \in F_{p^n}^*$. Then the equation $x^d = a$ over the finite field F_{p^n} has solutions if and only if $a^{(p^n-1)/l} = 1$, where $l = gcd(d, p^n - 1)$. Furthermore, if $x^d = a$ has solutions, then it has exactly l solutions.*

The degree of a vertex v in a graph G is the number of edges of G incident with v. We denote by $\Delta(G)$ the maximum degree of vertices of G. Now we can state our main results.

Theorem 2. *Let G be a graph with r vertices and t edges. Suppose the number of perfect matchings of G is $\phi(G)$. Let p be a prime, d be a positive integer and $a \in F_{p^n}^*$. Let $N(S(G, d, p^n, a))$ be the number of solutions of the associated system of equations $S(G, d, p^n, a)$. If $a^{(p^n-1)/l} = 1$, where $l = gcd(d, p^n - 1)$, then*

$$N(S(G, d, p^n, a)) \geq l^{r/2}\phi(G), \tag{1}$$

especially,

$$N(S(G, d, p^n, 1)) \geq l^{r/2}\phi(G), \tag{2}$$

furthermore, if $\Delta(G) \leq p$, then

$$N(S(G, p^n - 1, p^n, 1)) = (p^n - 1)^{r/2}\phi(G). \tag{3}$$

Proof. Since $a^{(p^n-1)/l} = 1$, by Lemma 1 the equation $x^d = a$ has exactly $l = gcd(d, p^n - 1)$ solutions. Let $E = \{b_1, b_2, \cdots, b_l\}$ be the set of these solutions.

Assume that the edge set of G is $\{e_1, e_2, \cdots, e_t\}$. Since G has $\phi(G)$ perfect matchings, let $\{F_1, F_2, \cdots, F_{\phi(G)}\}$ be the set of these perfect matchings.

For each $F_i \in \{F_1, F_2, \cdots, F_{\phi(G)}\}$, we can construct $l^{r/2}$ solutions of the associated system of equations $S(G, d, p^n, a)$ as follows: assume that $F_i = \{e_{i_1}, e_{i_2}, \cdots, e_{i_{r/2}}\}$ is a perfect matching of G, where $\{i_1, i_2, \cdots, i_{r/2}\} \subseteq \{1, 2, \cdots, t\}$. Let $X=(x_1, x_2, \cdots, x_t)^\top$ be a vector in $(F_{p^n})^t$ such that $x_j = c$ (where $c \in E$) if $j \in \{i_1, i_2, \cdots, i_{r/2}\}$, or $x_j = 0$ if $j \notin \{i_1, i_2, \cdots, i_{r/2}\}$.

By the definition of the associated system of equations, it is easy to check that X is a solution of $S(G, d, p^n, a)$. Note the set E (resp. F_i) has l (resp. $r/2$) elements, and G has $\phi(G)$ perfect matchings, so we can construct $l^{r/2}\phi(G)$ solutions of this kind like X, therefore $N(S(G, d, p^n, a)) \geq l^{r/2}\phi(G)$, this proves the inequality (1).

Especially, for the associated system of equations $S(G, d, p^n, 1)$, it is evident that the equation $x^d = 1$ has a solution $x = 1$. By Lemma 1, we know that the equation $x^d = 1$ has exactly $l = gcd(d, p^n - 1)$ solutions. By a similar discussion as above, we can also get that $N(S(G, d, p^n, 1)) \geq l^{r/2}\phi(G)$, this proves the inequality (2).

Now assume that $\Delta(G) \leq p$. For the associated system of equations $S(G, p^n - 1, p^n, 1)$, notice that $l = gcd(p^n - 1, p^n - 1) = p^n - 1$.

By above method, we can construct $(p^n - 1)^{r/2}\phi(G)$ solutions, let R be the set of these solutions. Let $X{=}(b_1, b_2, \cdots, b_t)^{\top}$ be an arbitrary solution of $S(G, p^n - 1, p^n, 1)$, in the following we will show that $\mathbf{X} \in R$.

Note that for any $b \in F^*_{p^n}$, by Lemma 1, b is always a solution of the equation $x^{p^n - 1} = 1$. Since the characterization of the finite field F_{p^n} is p and $\Delta(G) \leq p$. It implies that if the system of equations $S(G, p^n - 1, p^n, 1)$ has solution, then each equation in $S(G, p^n - 1, p^n, 1)$ has exactly one variable can take value in $F^*_{p^n}$. By the definition of the associated system of equations, $S(G, p^n - 1, p^n, 1)$ has exactly r equations, and each variable appears in exactly two equations (since each edge in G is incident with exactly two vertices). Thus X has exactly $r/2$ non-zero entries. Let $x_{j_1}, x_{j_2}, \cdots, x_{j_{r/2}}$ be those non-zero entries, where $\{j_1, j_2, \cdots, j_{r/2}\} \subseteq \{1, 2, \cdots, t\}$. The above discussion show that $\{e_{j_1}, e_{j_2}, \cdots, e_{j_{r/2}}\}$ is a perfect matching of G, and thus $\mathbf{X} \in R$. So $N(S(G, p^n - 1, p^n, 1)) = (p^n - 1)^{r/2}\phi(G)$, this proves the equality (3).

In algorithm complexity theory, the following result is well-known.

Lemma 3 ([2, 6]). *The problem of determining the number of perfect matchings in a graph is NP-hard.*

By above results, we can prove another main result of this paper.

Theorem 4. *The problem of determining the number of solutions to a system of equations over finite field is NP-hard.*

Proof. From the proof the equality (3) of Theorem 2, it is easy to see that the problem of determining the number of perfect matchings in a graph G can be reduced to the problem of counting the number of solutions of a special associated system of equations $S(G, p^n - 1, p^n, 1)$ over finite field F_{p^n}. By Lemma 3, we conclude that the problem of determining the number of solutions over finite field is NP-hard.

3 Conclusion

Estimating the number of solutions of equations over a finite field is an important, but difficult problem in coding theory. By Theorem 4, we also know that the problem is difficult in sense of complexity of algorithm.

Acknowledgements. This project is supported by the Natural Science Foundation of Fujian Province of China (No. 2010J01008) and the Science Foundation of Jimei University.

References

1. Bondy, J.A., Murty, U.S.R.: Graph theory with applications. Macmillan Press, London (1976)
2. Garey, M.R., Johnson, D.S.: Computers and intractability: A guide to the theory of NP-completeness. W.H. Freeman Company, San Francisco (1979)
3. Ireland, K., Rosen, M.: A classical introduction to modern number theory. Springer, New York (1982)
4. Katz, N.: An overview of deligine's proof of the riemann hypothesis for varieties over finite fields. In: Proc. Symp. Pure Math. Amer. Math. Soc., vol. 28, pp. 275–305 (2006)
5. Lidl, R., Niederreiter, H.: Finite fields, Encyclopedia of Mathematics and Its Applications. Addison-Wesley, Reading (1983)
6. Lovász, L., Plummer, M.D.: Matching theory. Elsevier Science, North-Holland, Amsterdam (1986)
7. Ren, D.B., Sun, Q., Yuan, P.Z.: Number of zeros of diagonal polynomials over finite fields. Finite Fields and Their Application 7, 197–204 (2001)
8. Weil, A.: Numbers of solutions of equations over finite fields. Bull. Amer. Math Soc. 55, 497–508 (1949)
9. Wolfmann, J.: Some systems of diagonal equations over finite fields. Finite Fields and Their Applications 4, 29–37 (1998)

The Stability of BAM Networks with Delayed Self-feedback and Impulses

Chao Chen and Zhen-kun Huang

School of Sciences, Jimei University, Xiamen, Fujian, 361021, China
e-mail: firedoctor@163.com, hzk974226@jmu.edu.cn

Abstract. This paper considers a class of bidirectional associative memory(BAM) neural networks with self-feedback and nonlinear impulses. A new criterion concerning global exponential stability for these neural networks is derived, which improve the previously known results. Finally, an example is given to show the applicability of our results.

Keywords: Bidirectional associative memory, neural networks, nonlinear impulses, delay.

1 Introduction

The bidirectional associative memory (BAM) model of the type

$$\begin{cases} \dot{x}_i(t) = -a_i x_i(t) + \sum_{j=1}^{n} p_{ij} f_j(y_j(t)) + r_i \\ \dot{y}_j(t) = -b_j y_j(t) + \sum_{i=1}^{n} q_{ji} g_i(x_i(t)) + s_j \end{cases} \tag{1}$$

known as an extension of the unidirectional autoassociator of Hopfield [1], was first introduced by Kosto [2]. It has been used in many fields such as optimization, pattern recognition and automatic control [3-6]. Realizing the ubiquitous existence of delay in neural networks, Gopalsamy and He [4] incorporated time delays into the model and considered the following system of delay differential equations

$$\begin{cases} \dot{x}_i(t) = -a_i x_i(t) + \sum_{j=1}^{m} c_{ij} f_j(y_j(t-\tau_{ij})) + r_i \ , \\ \dot{y}_j(t) = -b_j y_j(t) + \sum_{i=1}^{n} d_{ji} g_i(x_i(t-\sigma_{ij})) + s_j . \end{cases} \tag{2}$$

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 761–769.
springerlink.com

Moreover, it should be noted that the theory of impulsive differential equations is now being recognized to be not only richer than the corresponding theory of differential equations without impulse, but also represent a more natural framework for mathematical modelling of many real-world phenomena. there have been extensive results on the problem of the stability and other dynamical behaviors of impulsive BAM neural networks with delays in the literature, see [10-13] and the references cited therein. However, few authors have considered global exponential stability of BAM neural networks with nonlinear impulses which include common impulsive and non-impulsive system.

On the other hand, recent work [7-9] has shown that inhibitory self-connections play a role in stabilizing a network under some conditions on delays. In the present paper, inspired by references [7-9, 11, 13, 16-17], we shall study the stability problem of BAM neural networks with nonlinear impulsive and continuously distributed delays.

2 Preliminaried

In the following, we shall consider the nonlinear impulsive BAM neural networks with continuously distributed delays which can be described by the following integro-differential equations of the form

$$\begin{cases} \dot{x}_i(t) = -\rho_i(x_i(t)) + c_{ii}g_i(x_i(t-e_{ii})) + \sum\limits_{j=1}^{n} p_{ij}f_j(y_j(t-\tau_{ij})) + r_i, \quad t \geq 0,\ t \neq t_k, \\ \Delta x_i(t_k) = I_i(x_i(t_k)), \quad i = 1, 2, \cdots, n, \quad k = 1, 2, \cdots, \\ \dot{y}_j(t) = -\varrho_j(y_j(t)) + d_{jj}f_j(y_j(t-h_{jj})) + \sum\limits_{i=1}^{n} q_{ji}g_i(x_i(t-\sigma_{ji})) + s_j, \quad t \geq 0,\ t \neq t_k, \\ \Delta y_j(t_k) = J_j(y_j(t_k)), \quad j = 1, 2, \cdots, n, \quad k = 1, 2, \cdots, \end{cases} \tag{3}$$

where x_i and y_j are the activations of the ith neurons and the jth neurons, respectively. c_{ii}, d_{jj}, p_{ij}, q_{ji} are the connection weight, and r_i and s_j denote the external inputs. g_i, f_j $(i, j = 1, 2, \cdots, n)$ are signal transmission functions. Here $\Delta x_i(t_k) = x_i(t_k+0) - x_i(t_k-0)$, $\Delta y_j(t_k) = y_j(t_k+0) - y_j(t_k-0)$ are the impulses at moments t_k and $0 < t_1 < t_2 < \cdots$ is an increasing sequence. The nonlinear impulsive functions $I_i, J_j : R \rightarrow R$ are assumed to be continuous which detail form would be given in section 4. And system (3) is supplemented with initial values given by

$$x_i(s) = \varphi_i(s), \quad s \in [-\tau, 0], \quad \tau = \max_{1 \leq i,j \leq n}\{\tau_{ij}\} \quad i = 1, 2, \cdots, n,$$
$$y_j(s) = \psi_j(s), \quad s \in [-\sigma, 0], \quad \sigma = \max_{1 \leq i,j \leq n}\{\sigma_{ji}\} \ \ j = 1, 2, \cdots, n,$$

where $\varphi_i(\cdot)$, $\psi_j(\cdot)$ are bounded and continuous on $[-\tau, 0]$, $[-\sigma, 0]$, respectively.

In general speaking, by a solution of (3) we mean $z(t) = (x_1(t), \cdots, x_n(t), y_1(t), \cdots, y_n(t))^T \in R^{2n}$ in which $x_i(\cdot), y_j(\cdot)$ is piecewise continuous on $(0, \beta)$ for some $\beta > 0$ such that $z(t_k \pm 0)$ exists and $z(\cdot)$ is differentiable on intervals of the form $(t_{k-1}, t_k) \subset (0, \beta)$ and satisfies (3); we assume that $z(t)$ is left continuous with $z(t_k) = z(t_k - 0)(k = 1, 2, \cdots)$. Throughout this paper, we assume that:

•(H$_1$) $c_{ii}, d_{jj}, p_{ij}, q_{ji}, r_i, s_j \in R,\ \tau_{ij}, \sigma_{ji} \in [0, \infty)$. And $\rho_i, \varrho_j : R \to R$ are differentiable and strictly monotone increasing, i.e., $a_i = \inf\limits_{x \in R}\{\dot{\rho}_i(x)\} > 0$ and $b_j = \inf\limits_{x \in R}\{\dot{\varrho}_j(x)\} > 0,\ i, j = 1, 2, \cdots, n.$

•(H$_2$) f_j and g_i are Lipschitz-continuous on R with Lipschitz constant β_j $(j = 1, \cdots, n)$ and $\alpha_i (i = 1, \cdots, n)$ respectively, that is,

$$|f_j(x) - f_j(y)| \le \beta_j |x - y|, \quad |g_i(x) - g_i(y)| \le \alpha_i |x - y|, \quad \forall\ x,\ y \in R.$$

Throughout the paper, for convenience, we introduce the following notations: we will use $z = (x_1, x_2, \cdots, x_n, y_1, \cdots, y_n)^T \in R^{2n}$ to denote a column vector, in which the symbol "T" denotes the transpose of a vector, $P \ge 0$ denotes nonnegative matrix.

Before starting the main results, firstly, we shall give some definitions and lemmas as follows:

Definition 1. *A real matrix $H = (h_{ij})_{n \times n}$ is said to be a nonsingular M-matrix, if H has the form*

$$H = \alpha E - P, \qquad \alpha > 0, \quad P \ge 0,$$

where $\alpha > \rho(P)$, $\rho(P)$ denotes the spectral radius of P.

Definition 2. *A constant vector $z^* = (x_1^*, x_2^*, \cdots, x_n^*, y_1^*, y_2^* \cdots, y_n^*)^T$ is said to be the equilibrium point of (3) if it satisfies*

$$\begin{cases} \rho_i(x_i^*) = c_{ii} g_i(x_i^*) + \sum\limits_{j=1}^{n} p_{ij} f_j(y_j^*) + r_i, \quad i = 1, \cdots, n, \\ \varrho_j(y_j^*) = d_{jj} f_j(y_j^*) + \sum\limits_{i=1}^{n} q_{ji} g_i(x_i^*) + s_j, \quad j = 1, \cdots, n, \end{cases} \tag{4}$$

where the impulsive jumps $I_i(x_i^*) = J_j(y_j^*) = 0\ \ (i, j = 1, 2, \cdots, n)$.

Definition 3. *The unique equilibrium $z^* = (x_1^*, x_2^*, \cdots, x_n^*, y_1^*, y_2^* \cdots, y_n^*)^T$ of (3) is said to be globally exponentially stable, if there exist $\lambda > 0$ and $M \ge 1$ such that for all $t \ge 0$,*

$$\Big\{\sum_{i=1}^{n} |x_i(t) - x_i^*| + \sum_{j=1}^{n} |y_j(t) - y_j^*|\Big\} \le M e^{-\lambda t}\Big\{\sum_{i=1}^{n} \|\varphi_i - x_i^*\| + \sum_{j=1}^{n} \|\psi_j - y_j^*\|\Big\}.$$

Lemma 1. *([14]) Let $H = (h_{ij})_{n\times n}$ is a matrix with non-positive off-diagonal elements. Then H is an M-matrix if and only if there exists a positive diagonal $D = diag(d_1, \cdots, d_n)$ such that HD is a strictly diagonally dominant with positive diagonal entries; that is*

$$h_{ii}d_i > \sum_{j\neq i} |h_{ij}|d_j, \qquad i = 1, 2, \cdots, n.$$

Lemma 2. *([15]) Let $H \geq 0$ be an $n \times n$ matrix and $\rho(H) < 1$. Then $(E_n - H)^{-1} \geq 0$, where $\rho(H)$ denotes the spectral radius of H.*

3 Existence and Uniqueness of Equilibrium

Theorem 1. *In addition to $(H_1)-(H_2)$, assume further that $\rho(F) < 1$, where*

$$F = L^{-1}HD, \quad H = \begin{pmatrix} 0_{n\times n} & A_{n\times n} \\ B_{n\times n} & 0_{n\times n} \end{pmatrix}, \quad D = diag(\alpha_1, \alpha_2, \cdots, \alpha_n, \beta_1, \beta_2, \cdots, \beta_n),$$

$$L = diag(\Delta_1, \Delta_2, \cdots, \Delta_n, \Delta_1, \Delta_2, \cdots, \Delta_n), \;\; \Delta_i = \min\{\frac{a_i}{2}, \frac{b_i}{2}\}, \;\; A = (|d_{jj}| + |p_{ij}|)_{n\times n},$$

$B = (|c_{ii}| + |q_{ji}|)_{n\times n}$, then there exists a unique equilibrium for (3).

Proof. Consider a mapping $\Phi : R^{2n} \to R^{2n}$ defined by

$$\begin{cases} \rho_i(\Phi_i(z)) = c_{ii}g_i(x_i) + \sum_{j=1}^{n} p_{ij}f_j(y_j) + r_i, \quad i = 1, 2, \cdots, n, \\ \varrho_j(\Phi_{n+j}(z)) = d_{jj}f_j(y_j) + \sum_{i=1}^{n} q_{ji}g_i(x_i) + s_j, \quad j = 1, 2, \cdots, n. \end{cases} \tag{5}$$

We show that $\Phi : R^{2n} \to R^{2n}$ possesses a unique fixed point in R^{2n}. In fact, for any $z = (x_1, x_2, \cdots, x_n, y_1, y_2, \cdots, y_n)^T \in R^{2n}$, $\bar{z} = (\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n, \bar{y}_1, \bar{y}_2 \cdots, \bar{y}_n)^T \in R^{2n}$, from (5), we have

$$|\Phi_i(z) - \Phi_i(\bar{z})| = \frac{|\rho_i(\Phi_i(z)) - \rho_i(\Phi_i(\bar{z}))|}{|\dot{\rho}_i(\xi_i)|} \leq \frac{\alpha_i}{a_i}|c_{ii}||x_i - \bar{x}_i| + \sum_{j=1}^{n} \frac{\beta_j}{a_i}|p_{ij}||y_j - \bar{y}_j|$$

$$|\Phi_{n+j}(z) - \Phi_{n+j}(\bar{z})| = \frac{|\varrho_j(\Phi_{n+j}(z)) - \varrho_j(\Phi_{n+j}(\bar{z}))|}{|\dot{\varrho}_j(\xi_{n+j})|} \leq \frac{\beta_j}{b_j}|d_{jj}||y_j - \bar{y}_j| + \sum_{i=1}^{n} \frac{\alpha_i}{b_j}|q_{ji}||x_i - \bar{x}_i|$$

where $(\xi_1, \xi_2, \cdots, \xi_{2n})^T$ lies between $\Phi(z)$ and $\Phi(\bar{z})$.
Then, we get

$$|\Phi(z) - \Phi(\bar{z})| = F\Big(|x_1 - \bar{x}_1|, |x_2 - \bar{x}_2|, \cdots, |x_n - \bar{x}_n|, |y_1 - \bar{y}_1|, |y_2 - \bar{y}_2|, \cdots, |y_n - \bar{y}_n|\Big)^T,$$

Let η be a positive integer. We obtain

$$|\Phi^\eta(z) - \phi^\eta(\bar{z})| \le F^\eta\Big(|z-\bar{z}|_1, |z-\bar{z}|_2, \cdots, |z-\bar{z}|_{2n}\Big)^T,$$

Since $\rho(F) < 1$, we have

$$\lim_{\eta\to+\infty} F^\eta = 0,$$

which implies that there exist a positive integer N and a positive constant $\gamma < 1$ such that

$$F^N = (l_{ij})_{(2n)\times(2n)}. \quad \sum_{j=1}^{2n} l_{ij} \le 1, \quad i = 1, 2, \cdots, 2n,$$

So, we get

$$\|\Phi^N(z) - \Phi^N(\bar{z})\| \le \gamma\|z - \bar{z}\|,$$

which implies that the mapping $\Phi : R^{2n} \to R^{2n}$ is a contracting mapping. By the fixed point theorem of Banach Space, Φ possess a unique fixed point z^* in R^{2n} which is a unique solution of (3). Then, (3) has exactly one equilibrium.

4 Exponential Stability of Equilibrium

Theorem 2. *Suppose that all the conditions of Theorem 1 hold, and* $z^* = (x_1^*, \cdots, x_n^*, y_1^*, \cdots, y_n^*)^T$ *is the unique equilibrium of (3). Furthermore, assume that the following conditions are satisfied*

$$\begin{cases} I_i(x_i(t_k)) = -\gamma_{ik}(x_i(t_k) - x_i^*) + \int_{t_{k-1}}^{t_k} M_i(s)(x_i(s) - x_i^*)ds, i = 1, 2, \cdots, n, k \in Z^+, \\ J_j(y_j(t_k)) = -\delta_{jk}(y_j(t_k) - y_j^*) + \int_{t_{k-1}}^{t_k} N_j(s)(y_j(s) - y_j^*)ds, j = 1, 2, \cdots, n, k \in Z^+. \end{cases} \tag{6}$$

Then, the unique equilibrium z^* *of (3) is globally exponentially stable.*

Proof. Let $z(t) = (x_1(t), x_2(t), \cdots, x_n(t), y_1(t), y_2(t), \cdots, y_n(t))^T$ be an arbitrary solution of (3) with initial value $\Psi = (\phi_1(t), \phi_2(t), \cdots, \phi_n(t), \psi_1(t), \psi_2(t), \cdots, \psi_n(t))^T$. Set $u_i(t) = x_i(t) - x_i^*,\ v_j(t) = y_j(t) - y_j^*$, (3) can be reduced to the following system:

$$\begin{cases} \dot{u}_i(t) = -h_i(u_i(t)) + c_{ii}G_i(u_i(t - e_{ii})) + \sum_{j=1}^{n} p_{ij}F_j(v_j(t - \tau_{ij})) \\ \\ \dot{v}_j(t) = -k_j(v_j(t)) + d_{jj}F_j(v_j(t - h_{jj})) + \sum_{i=1}^{n} q_{ji}G_i(u_i(t - \sigma_{ji})) \end{cases} \tag{7}$$

for $t > 0,\ t \ne t_k,\ k \in Z^+$. And for any $i, j = 1, \cdots, n,\ k \in Z^+$,

$$\begin{cases} |u_i(t_k+0)| \le |1-\gamma_{ik}||u_i(t_k)| + \int\limits_{t_{k-1}}^{t_k} |M_i(s)||u_i(s)|ds, \\ |v_j(t_k+0)| \le |1-\delta_{jk}||v_j(t_k)| + \int\limits_{t_{k-1}}^{t_k} |N_j(s)||v_j(s)|ds, \end{cases} \tag{8}$$

where

$$h_i(u_i(t)) = \rho(x_i(t)) - \rho(x_i^*), k_j(v_j(t)) = \varrho(y_j(t)) - \varrho(y_j^*),$$

$$F_j(v_j(t)) = f_j(y_j(t)) - f_j(y_j^*), G_i(u_i(t)) = g_i(x_i(t)) - g_i(x_i^*).$$

Since $\rho(F) < 1$, it follows from Lemma 1 and Lemma 2 that $E - (HDL^{-1})^T$ is a nonsingular M-matrix, and there exists a diagonal matrix $D = diag(\xi_1, \xi_2, \cdots, \xi_n, \eta_1, \eta_2, \cdots, \eta_n)$, such that

$$\begin{cases} (1-\alpha_i|c_{ii}|)\xi_i > \sum\limits_{j=1}^{n} \alpha_i\eta_j|q_{ji}|, \quad i = 1, 2, \cdots, n, \\ (1-\beta_j|d_{jj}|)\eta_j > \sum\limits_{i=1}^{n} \beta_j\xi_i\|p_{ij}|, \quad j = 1, 2, \cdots, n. \end{cases} \tag{9}$$

Let $R_i(\lambda_i), Q_j(\varepsilon_j)$ be defined by

$$\begin{cases} R_i(\lambda_i) = \xi_i(1-\lambda_i) - \xi_i\alpha_i|c_{ii}|e^{\lambda_i e_{ii}} - \sum\limits_{j=1}^{n} \eta_j\alpha_i|q_{ji}|e^{\lambda_i\sigma_{ji}}, \quad i = 1, 2, \cdots, n, \\ Q_j(\varepsilon_j) = \eta_j(1-\varepsilon_j) - \eta_j\beta_j|d_{jj}|e^{\varepsilon_j h_{jj}} - \sum\limits_{i=1}^{n} \xi_i\beta_j|p_{ij}|e^{\varepsilon_j\tau_{ij}}, \quad j = 1, 2, \cdots, n, \end{cases}$$

where $\lambda_i, \varepsilon_j \in [0, \infty)$. Since $E - (HDL^{-1})^T$ is an M-matrix, it follows from (9) that

$$\begin{cases} R_i(0) = \xi_i - \xi_i\alpha_i|c_{ii}| - \sum\limits_{j=1}^{n} \alpha_i\eta_j|q_{ji}| > 0, \quad i = 1, 2, \cdots, n, \\ Q_j(0) = \eta_j - \eta_j\beta_j|d_{jj}| - \sum\limits_{i=1}^{n} \beta_j\xi_i\|p_{ij}| > 0, \quad j = 1, 2, \cdots, n. \end{cases}$$

Since $R_i(\cdot), Q_j(\cdot)$ are continuous on $[0, \infty)$ and $R_i(\lambda_i), Q_j(\varepsilon_j) \to -\infty$ as $\lambda_i, \varepsilon_j \to +\infty$, there exist $\lambda_i^*, \varepsilon_j^* > 0$ such that $R_i(\lambda_i^*) = 0, Q_j(\varepsilon_j^*) = 0$ and $R_i(\lambda_i) > 0$ for $\lambda_i \in (0, \lambda_i^*)$, $Q_j(\varepsilon_j) > 0$ for $\varepsilon_j \in (0, \varepsilon_j^*)$. By choosing a posivie constant $\lambda = \min\{\lambda_1^*, ..., \lambda_n^*, \varepsilon_1^*, ..., \varepsilon_n^*\}$, we have

$$\begin{cases} \xi_i(1-\lambda) - \xi_i\alpha_i|c_{ii}|e^{\lambda e_{ii}} - \sum\limits_{j=1}^{n} \eta_j\alpha_i|q_{ji}|e^{\lambda\sigma_{ji}} > 0, \quad i = 1, 2, \cdots, n, \\ \eta_j(1-\lambda) - \eta_j\beta_j|d_{jj}|e^{\lambda h_{jj}} - \sum\limits_{i=1}^{n} \xi_i\beta_j|p_{ij}|e^{\lambda\tau_{ij}} > 0, \quad j = 1, 2, \cdots, n. \end{cases} \tag{10}$$

Now we define

$$\begin{cases} U_i(t) = e^{\lambda t}|u_i(t)|, & \text{for } \ i = 1, 2, \cdots, n, \\ V_j(t) = e^{\lambda t}|v_j(t)|, & \text{for } \ j = 1, 2, \cdots, n. \end{cases} \tag{11}$$

Consider the following Lyapunov functional

$$\begin{aligned} V(t) = & \sum_{i=1}^{n} \xi_i \Big\{ U_i(t) + |c_{ii}|\alpha_i e^{\lambda e_{ii}} \int_{t-e_{ii}}^{t} U_i(s)ds + \sum_{j=1}^{n} \beta_j |p_{ij}| e^{\lambda \tau_{ij}} \int_{t-\tau_{ij}}^{t} V_j(s)ds \Big\} \\ & + \sum_{j=1}^{n} \eta_j \Big\{ V_j(t) + |d_{jj}|\beta_j e^{\lambda h_{jj}} \int_{t-h_{jj}}^{t} V_j(s)ds + \sum_{i=1}^{n} \alpha_i |q_{ji}| \int_{t-\sigma_{ji}}^{t} U_i(s)ds \Big\} \end{aligned} \tag{12}$$

The derivative of $V(t)$ along the trajectories of (3) is obtained as

$$\begin{aligned} D^+V(t) \le & - \sum_{i=1}^{n} \Big\{ (1-\lambda)\xi_i - \alpha_i \xi_i |c_{ii}| e^{\lambda e_{ii}} - \sum_{j=1}^{n} \eta_j \alpha_i |q_{ji}| e^{\lambda \sigma_{ji}} \Big\} U_i(t) \\ & - \sum_{j=1}^{n} \Big\{ (1-\lambda)\eta_j - \beta_j \eta_j |d_{jj}| e^{\lambda h_{jj}} - \sum_{i=1}^{n} \xi_i \beta_j |p_{ij}| e^{\lambda \tau_{ij}} \Big\} V_j(t) \\ \le & \ 0 \end{aligned} \tag{13}$$

for $t > 0, \ t \neq t_k, \quad k \in Z^+$. Also,

$$\begin{aligned} V(t_k + 0) \le & (1+\varepsilon_k)V(t_k) + m \int_{t_{k-1}}^{t_k} e^{\lambda(t_k - s)} V(s)ds \\ \le & \prod_{h=1}^{k} \Big[(1+\varepsilon_k) + m \frac{e^{\lambda(t_k - t_{k-1})} - 1}{\lambda} \Big] V(0), \quad k \in Z^+, \end{aligned} \tag{14}$$

where $\varepsilon_k = \max\{\gamma_{ik}, \delta_{ik}\}$.

From (12), we obtain

$$\begin{aligned} V(0) = & \sum_{i=1}^{n} \xi_i \Big\{ U_i(0) + |c_{ii}|\alpha_i e^{\lambda e_{ii}} \int_{-e_{ii}}^{0} U_i(s)ds + \sum_{j=1}^{n} \beta_j |p_{ij}| e^{\lambda \tau_{ij}} \int_{-\tau_{ij}}^{0} V_j(s)ds \Big\} \\ & + \sum_{j=1}^{n} \eta_j \Big\{ V_j(0) + |d_{jj}|\beta_j e^{\lambda h_{jj}} \int_{-h_{jj}}^{0} V_j(s)ds + \sum_{i=1}^{n} \alpha_i |q_{ji}| e^{\lambda \sigma_{ji}} \int_{-\sigma_{ji}}^{0} U_i(s)ds \Big\} \\ \le & \sum_{i=1}^{n} \Big[\xi_i + \xi_i \alpha_i |c_{ii}| e^{\lambda e_{ii}} e_{ii} + \sum_{j=1}^{n} \eta_j \alpha_i |q_{ji}| e^{\lambda \sigma_{ji}} \sigma_{ji} \Big] \|\phi_i - x_i^*\| \\ & + \sum_{j=1}^{n} \Big[\eta_j + \eta_j \beta_j |d_{jj}| e^{\lambda h_{jj}} h_{jj} + \sum_{i=1}^{n} \xi_i \beta_j |p_{ij}| e^{\lambda \tau_{ij}} \tau_{ij} \Big] \|\psi_i - y_j^*\|. \end{aligned} \tag{15}$$

Denote

$$\begin{aligned}\gamma_i^1 &= \xi_i + \xi_i\alpha_i|c_{ii}|e^{\lambda e_{ii}}e_{ii} + \sum_{j=1}^{n}\eta_j\alpha_i|q_{ji}|e^{\lambda\sigma_{ji}}\sigma_{ji}, \gamma_j^2\\ &= \eta_j + \eta_j\beta_j|d_{jj}|e^{\lambda h_{jj}}h_{jj} + \sum_{i=1}^{n}\xi_i\beta_j|p_{ij}|e^{\lambda\tau_{ij}}\tau_{ij}.\end{aligned}$$

It follows from (15), we have

$$\begin{aligned}&\sum_{i=1}^{n}|x_i(t)-x_i^*| + \sum_{j=1}^{n}|y_j(t)-y_j^*|\\ &\le Me^{-\lambda t}\prod_{h=1}^{k}\Big[(1+\varepsilon_k) + m\frac{e^{\lambda(t_k-t_{k-1})}-1}{\lambda}\Big]\Big\{\sum_{i=1}^{n}\|\phi_i - x_i^*\| + \sum_{j=1}^{n}\|\psi_j - y_j^*\|\Big\},\end{aligned}$$

where $t > 0$ and $M = \max\{\gamma_i^1, \gamma_j^2\}$ $(i, j = 1, 2, \cdots, n)$.

5 An Illustrative Example

In this section, an example is presented to illustrate the feasibility our results. We consider the following BAM neural networks:

$$\begin{cases}\dot{x}_i(t) = -a_ix_i(t) - g_i(x_i(t-0.5)) + \sum_{j=1}^{2}p_{ij}f_j(y_j(t-2) + r_i, \quad i = 1, 2, \quad t \ge 0,\ t \ne t_k,\\ \Delta x_i(t_k) = -\gamma_{ik}(x_i(t_k) - 1) + \int_{t_{k-1}}^{t_k}\cos(s)(x_i(s) - 1)ds,\\ \dot{y}_j(t) = -b_jy_j(t) - f_j(y_j(t-0.5)) + \sum_{i=1}^{2}q_{ji}g_i(x_i(t-3)) + s_j, \quad j = 1, 2, \quad t \ge 0,\ t \ne t_k,\\ \Delta y_j(t_k) = -\delta_{jk}(y_j(t_k) - 1) + \int_{t_{k-1}}^{t_k}\sin(s)(y_j(s) - 1)ds,\end{cases} \tag{16}$$

where $g_i(x) = f_j(x) = \dfrac{1}{2}\big(|x+1| - |x-1|\big)$, $\gamma_{ik} = 1 + \sin(5+3k)$, $\delta_{jk} = 1 + \cos(3k)$, $k \in Z^+$,
$(a_1, a_2)^T = (2, 2)^T$,$(b_1, b_2)^T = (3, 3)^T$,$(r_1, r_2)^T = (4, 2)^T$,$(s_1, s_2)^T = (3, 5)^T$,

$$\begin{pmatrix}p_{11} & p_{12}\\ p_{21} & p_{22}\end{pmatrix} = \begin{pmatrix}1 & -2\\ -3 & 4\end{pmatrix}, \qquad \begin{pmatrix}q_{11} & q_{12}\\ q_{21} & q_{22}\end{pmatrix} = \begin{pmatrix}-5 & 6\\ 7 & -8\end{pmatrix}.$$

So, by simple computation, we can see that (16) satisfy the conditions of Theorem 1 and Theorem 2, thus, (16) has exactly one equilibrium $(1, 1, 1, 1)^T$ which is globally exponentially stable.

Acknowledgments. The project is supported by the Foundation for Young Professors of Jimei University, the Scientific Research Foundation of Jimei University and the Foundation for Talented Youth with Innovation in Science and Technology of Fujian Province (2009J05009).

References

1. Hopfield, J.J.: Neuron with graded response have collective computational properties like those of two-state neurons. Proc. Natl. Acad. Sci. 81, 3088–3092 (1984)
2. Kosto, B.: Bi-directional associative memories. IEEE Trans. Syst. Man Cybernet 18, 49–60 (1988)
3. Kosto, B.: Adaptive bi-directional associative memories. Appl. Opt. 26, 4947–4960 (1987)
4. Gopalsamy, K., et al.: Delay-independent stability in bi-directional associative memory networks. IEEE Trans. Neural Networks 5, 998–1002 (1994)
5. Li, Y.K.: Global exponential stability of BAM neural networks with delays and impulses. Chaos, Solitons and Fractals 24, 279–285 (2005)
6. Li, Y.T., Yang, C.B.: Global exponential stability analysis on impulsive of BAM neural networks with distributed delays. J. Math. Anal. Appl. 324, 1125–1139 (2006)
7. Driessche, P., et al.: Stabilization role of inhibitory self-connections in a delayed neural network. Physica D 150, 84–90 (2001)
8. Wang, L., Zou, X.F.: Stabilization role of self-inhibitory connections in Cohn-Grossberg neural networks with general activation functions. Diff. Eqns. Dyn. Syst. 9, 341–352 (2001)
9. Wang, L., Zou, X.F.: Stability and bifurcation of bidirectional associative memory neural networks with delayed self-feedback. International Journal of Bifuraction and Chaos 15, 2145–2159 (2005)
10. Yang, H., Chu, T.: LMI conditions for stability of neural networks with distributed delays. Chaos, Solitons and Fractals 34, 557–563 (2007)
11. Lou, X.Y., Cui, B.T.: Global asymptotic stability of delay BAM neural networks with impulses. Chaos, Solitons and Fractals 29, 1023–1031 (2006)
12. Zhao, H.Y.: Exponential stability and periodic oscillatory of bidirectional associative memory neural network involving delays. Neurocomputing 69, 424–448 (2006)
13. Qiu, F., Cui, B.T., Wu, W.: Global exponential stability of high order recurrent neural network with time-varing delays. Appl. Math. Modelling 33, 198–210 (2009)
14. Horn, R.A., Johnson, C.R.: Topics in matrix analysis. Cambridge University Press, UK (1991)
15. Lasalle, J.P.: The Stability of Dynamical System. SIAM, Philadelphia (1976)
16. Akca, H., Alassar, R., Corachav, V.: Continuous-time additive Hopfield-type networks with impulses. J. Math. Anal. Appl. 290, 436–451 (2004)
17. Akca, H., Berezansky, L., Braverman, E.: On linear integro-differential equations with integral impulsive conditions. Z. Anal. Anwendwngen 15, 709–727 (1996)

The Selection and Rational Evaluation of Surface Mining Technology System

Shan-yi Liu, Zi-ling Song, and Wei Tong

Resources and environment college, Liao Ning Technical University,
Fuxin, 123000, Liaonig, China
e-mail: shanyi8231@sohu.com, songziling-163@163.com,
tongwei@163.com

Abstract. The most important part of surface mining design is to select the rational technology system. On the basis of the classification and according to various applicable conditions, we should determine a corresponding mining technology system for all sorts of opencast coalfield in order to give guidances on the mining design and to provide theoretical bases on the selection of mining technology. In this paper, it discussed how to determine a rational evaluation about mining technology. First, we should determine an evaluation index system of mining technology system which is constituted with 8 indexes of grade one, 20 indexes of grade two, 69 indexes of grade three. Second, we quantified the indexes of grade three according to the applicable conditions of various mining technology and concluded a quantification evaluation system. We used vectors which were constituted with the aggregate of numeric area of 69 indexes of grade three to indicate the quantification evaluation system and then calculated the value of subordination-degree and the value of weight coefficients. At last we could get the value of adaptability-degree of various mining technology, the mining technology with maximum value of adaptability-degree is the rational mining technology system.

Keywords: Surface mining technology, fuzzy evaluation, adptability-degree, evaluation index, subordination-degree.

1 Introduction

In the surface mining design of development of opencast coalfield, the most important decisive technique is the mining technology, however, it is difficult to select a corresponding and rational mining technology for a specific opencast coalfield. There are two main reasons: first, there are many factors to influence the surface mining technology; second, there are not specific evaluation criterias for surface mining technology. Traditionally the method to select mining technology is as follows: according to the characteristics of the opencast coalfield, it is proposed several feasible technological programs and then we make a detailed mining design for each program, compare the economic benefits and select a more economical technological program. Although this method is effective and

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 771–778.
springerlink.com

traditional, there is a phenomenon that the selected program may be not the optimal one because the optimal program may be missing when it is proposed. So the compared factors are not comprehensive. Especially our country lacks the research on the characteristics and the mining technology of nationwide opencast coalfield, so it leads to lack theoretical guidance and the forecasts of development trend of surface mining. At the same time, it also constrains the development strategies and the plan of long-term development[4-6].

The backward technology of mining machinery manufacturing also constrains the development of our surface mining, and they mutually restrictive and interactive. The way to solve the problem is formulating development strategies of surface mining and mining technology in order to promote the common progress of the industry of mining machinery manufacturing and surface mining. This paper attempts to discuss that through the classifiable research of national opencast coalfield, it proposes the applicable mining technology for different types of coalfields and makes a rational evaluation about it. Through establishing a comprehensive evaluation system and selecting a rational method of evaluation to evaluate the mining technology, it can provide the theoretical bases for the selection of a rational mining technology in the development of surface mining.

2 Classifiable System of Surface Mining Technology and Its Influencing Factors

Now there are many types of commonly used technology system. The applicable conditions for different types are different. As the lithology of coal seam, rock seam and the topsoil are quite different, the mining technology, rock stripping technology and topsoil stripping technology are usually different in large-scale opencast mine. In the mining design it is necessary to select mining technology, rock stripping technology and topsoil stripping technology separately. The commonly used stripping technologies are divided into: “independent discontinuous mining technology system: single dou excavator- auto mining technology” and other 23 kinds of technologies. The mining technologies are divided into: “semi-continuous mining technology system : single dou excavator- working face auto-semi-fixed crushing and screening station-belt conveyor mining technology” and other15 kinds of technologies.

For the opencast mine the influencing factors of selection mining technology include the natural conditions of opencast coalfield, the equipments for mining, mining design and so on. The following are specific factors:

(1) The natural conditions of opencast mine which include the nature of ore and rock, the burying conditions and burying depth of ore bodies, the slant of seam, thickness, layers, terrain, geographical location; climatic conditions; hydrology and engineering geology; and the types of opencast mine.

(2) Available equipments which include the following requirements: first, the equipments for mining, transportation and unloading should match the types and specifications; second, it is better to choose the sets of equipments which are unification in the same link; third, the provision of the equipment parts and the

maintance should be reliable; fourth, auxiliary equipments should be consistent with the main equipments; fifth, the training to users should be guaranteed; sixth, the sources of available equipments are reliable.

(3) Equipments can meet the requirements of production scale and the quality of ore. It is required to complete opencast stripping and coal mining operation and production. At the same time the selected equipments should be beneficial for the selecting and mining of coal, should meet the requirements of the quality of coal and should reduce the losses of coal.

(4) Considering the sources of funding to purchase the equipments. Because the equipments are expensive, the initial investerment for equipments is huge, it should be fully considered the sources and reliability of funding.

(5) Considering the requirements of environmental protection. Surface mining can ruin the environment seriously, so it is better to choose the technology which is less polluted on the environment and is beneficial to improving the environment.

(6) Considering the economic and technological conditions around the mines area. It is mainly considered the economic and technological level around coalfield, the conditions of transportation and the distance from urban centers, the conditions of energy supply and living conditions.

3 Rational Evaluation System of Surface Mining Technology

A. The Determination of the Evaluation Index System

According to the influencing factors of surface mining, the suitable conditions of various mining technology system as well as the characteristics of surface coalfield, we can conclude that an evaluation index system of surface mining is a three-level index system which includes eight indexes of grade one {A_1, A_2, A_3, A_4, A_5, A_6, A_7, A_8} ; 20 indexes of grade two {A_{11}, A_{12}, A_{13}, A_{14}, A_{15}}, {A21、A22}, {A31, A32}, {A41, A42}, {A51, A52}, {A61, A62, A63, A64, A65}, {A71, A72}, {A81, A82} ; 69 indexes of grade three {x1, x2, x3, x4}, {x5, x6, x7, x8}, ……, {x68、x69}.We can see the significance of various indexes from the directions in the 4th, 2008 “the learned journal of coal”. The former six indexes （A1, A2, A3, A4, A5, A6） of grade one and the 8th th index（A8） of grade one are the same as the classifiable index system of coalfield. The constitution of the 7th index（A7）of grade one, its index of grade two and its index of grade three is shown in the following fig.1:

B. The Selection of Evaluation Methods and the Establishment of Models

Through the fuzzy comprehensive evaluation of the above mentioned 69 indexes of grade three, we get an adaptability index of different mining technology （B_i，C_i） for a opencast coalfield. We call it adaptability-degree. The technology with the maximum value of adaptability-degree is the suitable mining technology.

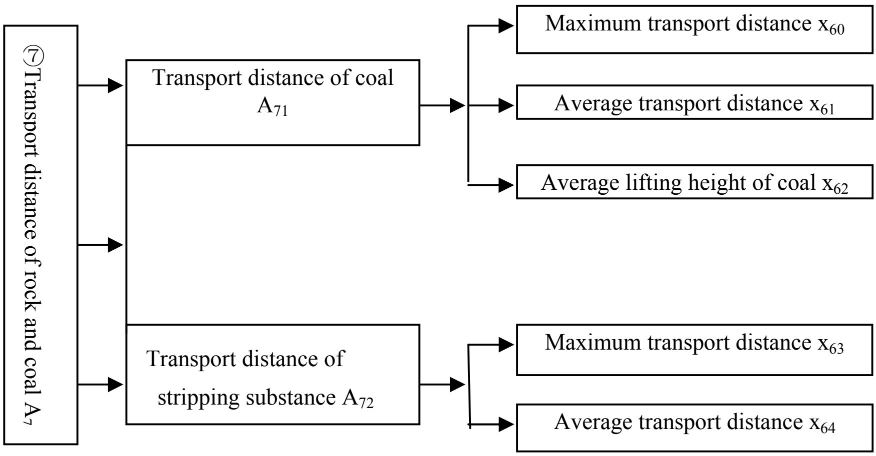


Fig. 1. The constitution of index of grade one A_7

The procedures and methods of establishment of the evaluation model are as follows:

Step 1. The determination of applicable conditions for all sorts of mining technology system

According to relevant references [1] [2] [3] and the experience of surface mining design, the applicable conditions of all sorts of mining technology system are as follows:

(1) Independent discontinuous mining technology system: single dou (one dou equals to 10 liters) excavator- automotive mining technology system（$B_{1,}$ C1）is suitable for the following conditions :① The open-air coal mine whose the terrain and ore bodies are complex and the length is limited;② The transport distance is not more than 3～5km; ③The opencast mine whose stripping volume is large and the construction building is fast; ④ Deep pioneering transportation; ⑤ For any lithology.

(2) Independent discontinuous mining technology system: single dou (one dou equals to 10 liters) excavator- railway mining technology system（$B_{2,}$ C2）is suitable for the following conditions: ① All types of ore and rock and bulk of materials; ② For any climate; ③ As the cost of railway transport is less, it is suitable for the coal mine which is large capacity and the long transport distance; ④ Because of the poor climbing ability of railway transport（20~40%） and the restriction on the restricted gradient, the mining size and mining depth are limited. Usually, for the standard gauge railway transport, the lengthy of bottom boundary is not less than 1.2Km; the height of sloping open-pit mine should be around 200m and the height of hollow opencast coal mine should be around 100~200m,

sometimes it can reach 200~300m ; ⑤ The expected life should be long enough to pay the higher capital investment.

(3) Independent continuous mining technology: wheel dou (one dou equals to 10 liters) excavator—belt conveyor mining technology system （$B_{7,}$ C_7） is suitable for the following conditions: ① The materials whose hardness is f=1~2 from economic aspect and the cutting resistance $K_L \leq 100kg/cm$, $K_F \leq 6{\sim}7kg/cm^2$; ② The climate is not too cold; ③ Tthe seam memory is more regular; ④ Materials don’t contain abrasive materials or easily clogging materials.

(4) Combined discontinuous mining technology system: drag—shoveling, stripping, inverted heaping throwing—blasting mining technology system（B_{10}） is suitable for the following conditions: ① The coal seam should be horizontal or similar to horizontal or little tilting inn order to ensure the adequate space and stability of dump; ②The thickness of inverted heaping and stripping materials is not too thick; ③ The coal seam is not too thick; ④The stripping materials are medium hard rock or hard rock which have effective blasting (the block is uniform, and the bulk is less); ⑤ For any climate.

(5) Combined continuous mining technology system: wheel-dou (one dou equals to 10 liters) excavator—transport dumping bridge mining technology system（B_{14}） is suitable for the following conditions: ① The stripping materials are loose and soft but there is a certain carrying capacity; ② The coal seam should be horizontal or similar to horizontal; ③ There is no major fault structure; ④ There are large coal reserves and the expected life of pen-air mine is long; ⑤The climate is not too cold.

(6) Semi-continuous mining technology system: wheel-dou (one dou equals to 10 liters) excavator-automotive (or railway) mining technology （$B_{19,}$ C_{11}） is suitable for the following conditions: ① The open-air coal mine which strips the loose and soft rock and in the situation that the belt conveyor is not suitable or economic; ② Working face is not horizontal and straight.

(7) Semi-continuous mining technology system: single dou (one dou equals to 10 liters) excavator- working face auto- semi fixed crushing screening station—belt conveyor mining technology system（$B_{21,}$ C_{13}） is suitable for the following conditions: ① The mining of coal seam; ② The stripping of hard and medium hard rock and effective blasting; ③ Long-distance transport and high promoting degree; ④ The climate is not too cold.

There are the other 16 types of mining technology systems such as independent discontinuous mining technology system: hydraulic excavator automotive mining technology system, independent continuous mining technolgoy: chain dou (one dou equals to 10 liters) excavatour-belt conveyor mining technology system, combined continuous mining technology system: wheel-dou (one dou equals to 10 liters) with dumping cantilever excavator mining technology system and semi-continuous mining technology system: single-dou excavator-mobile working face crusher-belt conveyor mining technology system. Because we were familiar with the above 16 types of mining technology systems, we didn’t discuss it any more in this paper.

Step 2. Processing the above applicable conditions of mining technology to quantify according to the evaluation indexes (x_1～x_{69}) of grade three in technology evaluation system ,we can get a quantification evaluation system for various applicable conditions of technology.

Using the vector: $U = \{[u_{1N}, u_{1M}], [u_{2N}, u_{2M}], [u_{3N}, u_{3M}], \cdots\cdots, [u_{69N}, u_{69M}]\}$ to indicate quantification index system of mining technology and the $[u_{kN}, u_{kM}]$ is the numeric area of the index $\mathbf{x_k}$ of grade three which is concluded from one of applicable conditions of technology. We standardized the numeric area $[u_{kN}, u_{kM}]$ of various evaluation index $\mathbf{x_k}$ according to the formula in reference [7]. At last we get the quantification index vectors of applicable conditions of various mining technology system (in fact, it is the aggregate of numeric area of 69 indexes of grade three), as follows

$$U_{Bi} = \{[u_{1N}, u_{1M}]_{Bi}, [u_{2N}, u_{2M}]_{Bi}, \cdots\cdots, [u_{69N}, u_{69M}]_{Bi}\}, \tag{1}$$

i=1,2,3,………,23, indicating 1～23 stripping mining technologies.

$$U_{Ci} = \{[u_{1N}, u_{1M}]_{Ci}, [u_{2N}, u_{2M}]_{Ci}, \cdots\cdots, [u_{69N}, u_{69M}]_{Ci}\}, \tag{2}$$

i=1,2,3,………,15, indicating 1～15 mining technologies.

Step 3. Calculating the subordination-degree $\mu_{Bi}(x_k), \mu_{Ci}(x_k)$ for relative mining technology of every evaluation index x_k for the evaluated opencast coalfield:

$$\mu_{Bi}(x_k) = \begin{cases} 1, & when \;\; x_k \subset [u_{kN}, u_{kM}]_{Bi} \\ 0, & when \;\; x_k \not\subset [u_{kN}, u_{kM}]_{Bi} \end{cases} \tag{3}$$

$$\mu_{Ci}(x_k) = \begin{cases} 1, & when \;\; x_k \subset [u_{kN}, u_{kM}]_{Ci} \\ 0, & when \;\; x_k \not\subset [u_{kN}, u_{kM}]_{Ci} \end{cases} \tag{4}$$

If the value of x_k belongs to the numeric area [x_{Kn}, x_{kM}], so the calculating formula of relative subordination-degree is as follows:

$$\mu_{Bi}(x_k) = \begin{cases} 1, & when \quad [x_{kN}, x_{kM}] \subset [u_{kN}, u_{kM}]_{Bi} \\ 0, & when \quad [x_{kN}, x_{kM}] \not\subset [u_{kN}, u_{kM}]_{Bi} \\ \dfrac{u_{kM} - x_{kN}}{x_{kM} - x_{kN}}, & when \quad x_{kN} \subset [u_{kN}, u_{kM}]_{Bi}, \quad x_{kM} \not\subset [u_{kN}, u_{kM}]_{Bi} \\ \dfrac{x_{kM} - u_{kN}}{x_{kM} - x_{kN}}, & when \quad x_{kN} \not\subset [u_{kN}, u_{kM}]_{Bi}, \quad x_{kM} \subset [u_{kN}, u_{kM}]_{Bi} \end{cases} \tag{5}$$

By equation(5), it can also be calculated the value of $\mu_{Ci}(x_k)$ of mining technology. k=1,2,3,……,69 indicating 69 indexes of grade three.

Step 4. According to the refernce[1], we can determine the weight coefficients w_k of every evaluation index.

The vector of weight coefficient: $W = (w_1, w_2, \cdots\cdots, w_{69})$

Step 5. Calculating the value of fitness-degree R_{Bi}, R_{Ci} of every mining technology.

$$R_{Bi} = \sum_{k=1}^{69} \mu_{Bi}(x_k) \cdot w_k \ , \tag{6}$$

$$R_{Ci} = \sum_{k=1}^{69} \mu_{Ci}(x_k) \cdot w_k \ . \tag{7}$$

Step 6. The determination of rational mining technology

According to the value of adaptability-degree R_{Bi}, R_{Ci} of various mining technology, the mining technology with maximum value of adaptability-degree is the rational mining technology.

Stripping technology: $B_I = Max\{R_{B1}, R_{B2}, R_{B3}, \cdots\cdots\}$;

Mining technology: $C_I = Max\{R_{C1}, R_{C2}, R_{C3}, \cdots\cdots\}$.

If we do a adaptability evaluation on a mining technology system which is selected by an opencast coalfield, we can get the index values of fitness-degree R_B, R_C, the value of fitness-degree is the number in the area [0, 1], the criteria of the adaptability-degree is as follows:

$R_B \subset [0.9, 1.0]$, indicating the technology is entirely suitable for the opencast coalfield.
$R_B \subset [0.8, 0.9)$, indicating the technology is suitable for the opencast coalfield.
$R_B \subset [0.7, 0.8)$, indicating the technology is less suitable for the opencast coalfield.
$R_B \subset [0.6, 0.7)$, indicating the technology is not very suitable for the opencast coalfield.
$R_B \subset [0.0, 0.6)$, indicating the technology is not suitable for the opencast coalfield.

4 Conclusion

(1) The factors that influence the selection of the surface-mining technology system including: the natural conditions of mineral deposit, available equipments, production scale, requirements for ore quality, sources of funding for purchasing equipment, requirements of environmental protection, economic and technological conditions around mining area and so on;

(2) It makes a systematic classification of the surface-mining technology, it determines 23 types of commonly used strip mining technology and 15 mining technology; it also identifies a variety of mining technology in general use;
(3) It establishes a evaluation index system for the adaptability evaluation of surface mining technology system, including 8 indexes of grade one, 20 indexes of grade two, 69 indexes of grade three;
(4) It determines the adaptability evaluation methods of mining technology, it quantifies various applicable conditions of surface mining technology and it establishes adaptability subordination-degree function between surface coalfield and mining technology. Through calculating the value of subordination-degree of every factor and calculating comprehensive adaptability value R_{Bi}, R_{Ci} according to weight coefficients of every factor, it concludes that the type of mining technology with the maximum value of adaptability-degree is the rational mining technology system.

References

1. Zi-ling, S.: The Research of Modern Design Theory and Methods of Opencast Mine Liaoning Technical University (2007)
2. Zhong-zhou, L.: The Study of Surface Ming, vol. 1. China Ming University (1990)
3. Rong-xin, Y.: The Study of Surface Ming, vol. 2. China Ming University (1990)
4. China Ming University editing, Handbook of Surface Mining. Coal Industry, Peking (1987)
5. Brady, T.M.: Continue mining technology system of surface mine ‘Bulk Handling in open pit mines & quarries’, pp. 131–138 (1986)
6. Rushou, D.: New Development of Design Theory for Opencast Mine. The Learned Journal of China Ming University 32(1), 1–7
7. Zi-ling, S., Yun-dong, M.: The Classifiable Research on Our National Opencast Coalfield (I) — the Establishment of Classifiable Index. The Learned Journal of Coal 9, 1002–1005 (2008)

The Blow-Up of Discrete Solution for NLS Equation with Potential

Zong-qi Liang

School of Science, Jimei University, Xiamen 361021 P.R. China
e-mail: zqliang@jmu.edu.cn

Abstract. In this paper we consider the blow-up for initial and Dirichlet boundary-value problem of a class of nonlinear Schrödinger equations with potential. We establish a conservative difference spectral approximation and propose sufficient conditions for the blow-up of approximation as well as the maximum time interval of existence of the solution. Finally it is proved that the spatial location of the radially symmetric blow-up solution is the origin.

Keywords: Nonlinear schrödinger equation, blow-up, spectral method.

1 Introduction

Nonlinear Schrödinger equation is a fundamental model in quantum mathematics mechanics. Classical nonlinear Schrödinger equation (without potential) is used to describe some phenomena is quantum physics, such as the propagation of laser beam in dispersive and nonlinear medium, self trapping in nonlinear optics ([1]) and Langmur waves in plasma ([2]). The nonlinear Schrödinger equation with potential has also definite physical background, especially the nonlinear Schrödinger equation with a harmonic potential is known as a model for describing the remarkable Bose-Einstein condensate (BEC) ([3,4]).

In this paper we consider the following the nonlinear Schrödinger equation with potential

$$\begin{cases} iu_t + \triangle u - v(x)u + \lambda f(|u|^2)u = 0, & (x,t) \in \Omega \times [0,T), \quad (1.1) \\ u(x,0) = u_0(x), & x \in \overline{\Omega}, \quad (1.2) \\ u(x,t)|_{x\in\partial\Omega} = 0. & t \geq 0, \quad (1.3) \end{cases}$$

where $i = \sqrt{-1}, u_t = \frac{\partial u}{\partial t}, \Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \cdots + \frac{\partial^2}{\partial x_d^2}$, Ω is a bounded domain in R^d, with boundary $\partial\Omega$ and $\overline{\Omega} = \Omega \cup \partial\Omega$. $u(x,t)$ is a complex-valued function defined on $\overline{\Omega} \times [0,T)$, $v(x)$ is a known real function, and $\lambda \geq 0$ is a real

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 779–789.
springerlink.com

parameter, $f(s) = s^p$. ReZ denotes real part of Z, ImZ denotes imaginary part of Z.

The equation has infinitely many conservation laws, include L^2 norm, the energy and Momentum, whose explicit form is.

•L^2-norm:

$$\frac{d}{dt}E_0(u) = 0, \quad \text{where } E_0(u) = \int_\Omega |u|^2 dx. \tag{1.4}$$

• Energy:

$$\frac{d}{dt}E_1(u) = 0, \quad \text{where } E_1(u) = \int_\Omega |\nabla u|^2 dx + \int_\Omega v(x)|u|^2 dx - \frac{\lambda}{p+1}\int_\Omega |u|^{2p+2} dx. \tag{1.5}$$

The nonlinear Schrödinger equation (1.1) is a typical dispersive wave equation, which reflects the relation between dispersion and nonlinear interaction. When dispersion dominate, energy disperse in space and solution exists globally, decaying with time evolving ([5,6]). When the dispersion and nonlinear reach balance, the nonlinear Schrödinger equation has localized, finite energy solutions which are often standing waves ([7]). When the nonlinearity dominate, wave will collapse and the solution blow up in finite time ([8]).

In this paper we are interested in the numerical approximation of blow-up the initial and Dirichlet boundary-value problem for a class of nonlinear Schrödinger equations with potential. Many numerical schemes are used to simulate the nonlinear Schrödinger Equation [9] including finite difference[10], finite element [11] and pseudo-spectral [12] schemes. Notice that the paper about numerical approximation for blow-up of nonlinear Schrödinger equation are few considered. In [13,14] Akrivis, Dougalis [13] and Salvador Jiménez [14] only studied that numerical approximation in radially symmetric case for $v(x) = 0, \lambda = 1, p = 2$.

The organization of the paper is as follows. In Section 2, we established a fully discrete spectral scheme, that is, which carry out discrete in time. In section 3, it is proven that the approximate solutions satisfy three conservation laws, a sufficient condition for the blow-up approximate solution is given as well, i.e. the solution would blow up if the initial value with certain conditions, an explicit upper bound of the T interval of existence of the solution is obtained.

2 Fully Discrete Spectral Method

For $1 \le q \le \infty$, we introduce the space $L^q(\Omega) = \{v;\ \|v\|_{L^q} < \infty\}$, where the corresponding norm is denoted by

$$\|v\|_{L^q} = \begin{cases} \left(\int_\Omega |v|^q dx\right)^{\frac{1}{q}}, 1 \le q < \infty, \\ ess \sup_{x\in\Omega} |v(x)|, \qquad q = \infty. \end{cases}$$

We recall that $(\cdot,\cdot)$ and $\|\cdot\|$ are the inner products and the norms of L^2 for $p = 2$ respectively, where $(u, v) = \int_\Omega u(x)\overline{v(x)}dx$, $\overline{v}$ is the conjugate function of v.

Let $\partial_x^k v = \frac{\partial^k v}{\partial x^k}$ for any positive integer m. We define the following space $H^m(\Omega)$

$$H^m(\Omega) = \{v;\ \partial_x^k v \in L^2(\Omega),\ 0 \le k \le m\}.$$

For any fixed integer$N \ge 1$, assume that $P_N(\Omega)$ is the space of all algebraic polynomials with complex coefficients of degree, i.e., the degree of each $x_j (j = 1, 2, \cdots, d)$. The piecewise polynomial space is defined as follows: $P_N(\Omega) = \{\phi \in L^2(\Omega),\ \phi|_\Omega \in P_N(\Omega)\}$, $V_N = H^1(\Omega) \cap P_N(\Omega), V_N^0 = H_0^1(\Omega) \cap P_N(\Omega)$.

Let $\Delta t = \frac{T}{M}$ is the time step and $t_n = n\Delta t(n = 0, 1, 2, \cdots, M)$, where M is a positive integer. We seek u^n approximating $u(t_n)$ and satisfying for $n = 0, 1, 2, \cdots, M - 1$.

Define

$$\phi^*(u, v) = \frac{1}{p+1}\frac{u^{p+1} - v^{p+1}}{u - v} = \frac{u^p + u^{p-1}v + \cdots + uv^{p-1} + v^p}{p+1},$$

$$\phi(u, v) = \phi^*(|u|^2, |v|^2).$$

We construct the following scheme of fully-discrete approximation: find $u_N^{n+1} \in V_N^0,\ n = 0, 1, \cdots, M - 1$ such that

$$\begin{cases} i(\frac{u_N^{n+1} - u_N^n}{\triangle t}, \chi) - (\frac{\nabla u_N^{n+1} + \nabla u_N^n}{2}, \nabla\chi) - (v(x)\frac{u_N^{n+1} + u_N^n}{2}, \chi) \\ \qquad + \lambda(\phi(u_N^{n+1}, u_N^n)\frac{u_N^{n+1} + u_N^n}{2}, \chi) = 0, \quad \forall\chi \in V_N^0, & (2.1) \\ u_N^0 = P_{1,N}^0 u_0. & (2.2) \end{cases}$$

3 Conservation Laws and Sufficient Conditions for Blow-Up of Approximate Solution

It is commonly accepted that to simulate Hamiltonian wave processes symmetric and conservative schemes are preferred over conventional ones because of their better global stability and long time behavior. Therefore we obtain the following theorem.

Theorem 3.1. *The solutions u_N^n $(\forall n = 0, 1, \cdots, M-1)$ of (2.1)-(2.2) satisfy the following conservation laws:*

$$E_0(u_N^n) = \|u_N^n\| = \|u_N^0\| = E_0(u_N^0), \tag{3.1}$$

$$E_1(u_N^n) = \|\nabla u_N^n\|^2 + \int_\Omega v(x)|u_N^n|^2 dx - \frac{\lambda}{p+1}\|u_N^n\|_{L^{2p+2}}^{2p+2} = E_1(u_N^0). \tag{3.2}$$

Proof. Let $\chi = u_N^n + u_N^{n+1}$ in (2.1), taking the imaginary part, we obtain the (3.1) easily. Similarly, (3.2) can be reduced by taking the real part of $\chi = u_N^n - u_N^{n+1}$.

For convenience, let $\varphi^n = \frac{1}{2}(u_N^n + u_N^{n+1})$, then the scheme (2.1) can be rewritten as

$$i(\frac{u_N^{n+1} - u_N^n}{\triangle t}, \chi) - (\nabla\varphi^n, \nabla\chi) - (v(x)\varphi^n, \chi) + \lambda(\phi(u_N^{n+1}, u_N^n)\varphi^n, \chi) = 0. \tag{3.3}$$

In the following, we consider about the blow-up of estimate solution.

Theorem 3.2. *Suppose that* $p \geq \frac{4}{d\lambda}, 2v(x) + x\cdot\nabla v(x) \geq 0$, *and if* $u_0(x) \in H_0^1(\Omega)$ *satisfies either of the following conditions,*

(1) $E_1(u_N^0) < 0$;
(2) $E_1(u_N^0) = 0, S_0 < 0$;
(3) $E_1(u_N^0) > 0, S_0 < 0, S_0^2 > 16E_1(u_N^0)W_0$;

where

$$S_n = S(t_n) = 2Im\int_\Omega [(x\cdot\nabla u_N^{n+1})\overline{u}_N^{n+1} + (x\cdot\nabla u_N^n)\overline{u}_N^n]dx, \tag{3.4}$$

and

$$W_n = W(t_n) = \int_\Omega |x|^2|u_N^n|^2 dx, \tag{3.5}$$

then the solution u_N^n *of (2.1)-(2.2) does not globally exist and the time interval is bounded by*

$$T \leq T^* = \frac{2W_0}{\sqrt{S_0^2 - 16E_1(u_N^0)W_0} - S_0}. \tag{3.6}$$

Proof. Let $\chi = |x|^2\varphi^n$ in equations (3.3), then taking imaginary part, we obtain

$$Re\int_\Omega \frac{u_N^{n+1} - u_N^n}{\Delta t}|x|^2(\overline{u}_N^{n+1} + \overline{u}_N^n)dx = 4Im\int_\Omega (x\cdot\nabla\varphi^n)\overline{\varphi}^n dx, \tag{3.7}$$

i.e.,

$$\frac{W_{n+1} - W_n}{\triangle t} = \frac{\int_\Omega |x|^2|u_N^{n+1}|^2dx - \int_\Omega |x|^2|u_N^n|^2dx}{\Delta t} = 4Im\int_\Omega (x\cdot\nabla\varphi^n)\overline{\varphi}^n dx. \tag{3.8}$$

Set $F_n = 4Im\int_\Omega (x\cdot\nabla\varphi^n)\overline{\varphi}^n dx$,

$$\begin{aligned}\frac{F_n - F_{n-1}}{\Delta t} &= \frac{4}{\Delta t}Im[\int_\Omega (x\cdot\nabla\varphi^n)\overline{\varphi}^n dx - \int_\Omega (x\cdot\nabla\varphi^{n-1})\overline{\varphi}^{n-1}dx]\\ &= -\frac{2d}{\Delta t}Im\int_\Omega (\overline{\varphi}^n + \overline{\varphi}^{n-1})(\varphi^n - \varphi^{n-1})dx\\ &\quad -\frac{4}{\Delta t}Im\int_\Omega [x\cdot(\nabla\overline{\varphi}^n + \nabla\overline{\varphi}^{n-1})](\varphi^n - \varphi^{n-1})dx\\ &= I_1 + I_2, \end{aligned} \tag{3.9}$$

in the following, we deduce the estimations of I_1 and I_2, respectively,

$$\begin{aligned} I_1 &= -\frac{2d}{\Delta t}Im\int_\Omega (\overline{\varphi}^n + \overline{\varphi}^{n-1})(\varphi^n - \varphi^{n-1})dx\\ &= -dIm[\int_\Omega \overline{\varphi}^n\frac{u_N^{n+1} - u_N^n}{\Delta t}dx + \int_\Omega \overline{\varphi}^{n-1}\frac{u_N^n - u_N^{n-1}}{\Delta t}dx\\ &\quad + \int_\Omega \overline{\varphi}^{n-1}\frac{u_N^{n+1} - u_N^n}{\Delta t}dx + \int_\Omega \overline{\varphi}^n\frac{u_N^n - u_N^{n-1}}{\Delta t}dx]\\ &= -2dIm\int_\Omega \overline{\varphi}^n\frac{u_N^{n+1} - u_N^n}{\Delta t}dx - 2dIm\int_\Omega \overline{\varphi}^{n-1}\frac{u_N^n - u_N^{n-1}}{\Delta t}dx\\ &\quad +2dIm\int_\Omega \frac{\overline{u}_N^n - \overline{u}_N^{n-1}}{2}\frac{u_N^{n+1} - u_N^n}{\Delta t}dx. \end{aligned} \tag{3.10}$$

$$\begin{aligned} I_2 &= -\frac{4}{\Delta t}Im\int_\Omega [x\cdot(\nabla\overline{\varphi}^n + \nabla\overline{\varphi}^{n-1})](\varphi^n - \varphi^{n-1})dx\\ &= -2Im[\int_\Omega (x\cdot\nabla\overline{\varphi}^n)\frac{u_N^{n+1} - u_N^n}{\Delta t}dx + \int_\Omega (x\cdot\nabla\overline{\varphi}^{n-1})\frac{u_N^n - u_N^{n-1}}{\Delta t}dx\\ &\quad + \int_\Omega (x\cdot\nabla\overline{\varphi}^{n-1})\frac{u_N^{n+1} - u_N^n}{\Delta t}dx + \int_\Omega (x\cdot\nabla\overline{\varphi}^n)\frac{u_N^n - u_N^{n-1}}{\Delta t}dx]\\ &= -4Im[\int_\Omega (x\cdot\nabla\overline{\varphi}^n)\frac{u_N^{n+1} - u_N^n}{\Delta t}dx + \int_\Omega (x\cdot\nabla\overline{\varphi}^{n-1})\frac{u_N^n - u_N^{n-1}}{\Delta t}dx]\\ &\quad -2Im\int_\Omega (x\cdot\frac{\nabla\overline{u}_N^n - \nabla\overline{u}_N^{n-1}}{2})\frac{u_N^n - u_N^{n-1}}{\triangle t}dx\\ &\quad +2Im\int_\Omega (x\cdot\frac{\nabla\overline{u}_N^{n+1} - \nabla\overline{u}_N^n}{2})\frac{u_N^{n+1} - u_N^n}{\triangle t}dx\\ &\quad -2dIm\int_\Omega \frac{\overline{u}_N^n - \overline{u}_N^{n-1}}{2}\frac{u_N^{n+1} - u_N^n}{\Delta t}dx. \end{aligned} \tag{3.11}$$

In terms of (3.10) and (3.11) can be rewritten as

$$\begin{aligned}
&\frac{F_n - F_{n-1}}{\Delta t} - 2Im\int_\Omega (x\cdot \frac{\nabla\overline{u}_N^{n+1} - \nabla\overline{u}_N^{n}}{2})\frac{u_N^{n+1}-u_N^n}{\triangle t}dx \\
&+2Im\int_\Omega (x\cdot \frac{\nabla\overline{u}_N^{n} - \nabla\overline{u}_N^{n-1}}{2})\frac{u_N^{n}-u_N^{n-1}}{\triangle t}dx \\
= &-2dIm\int_\Omega \overline{\varphi}^n\frac{u_N^{n+1}-u_N^n}{\Delta t}dx - 2dIm\int_\Omega \overline{\varphi}^{n-1}\frac{u_N^{n}-u_N^{n-1}}{\Delta t}dx \\
&-4Im\int_\Omega (x\cdot\nabla\overline{\varphi}^n)\frac{u_N^{n+1}-u_N^n}{\Delta t}dx \\
&-4Im\int_\Omega (x\cdot\nabla\overline{\varphi}^{n-1})\frac{u_N^{n}-u_N^{n-1}}{\Delta t}dx. \qquad (3.12)
\end{aligned}$$

Noting $S_n = 2Im\int_\Omega[(x\cdot\nabla u_N^{n+1})\overline{u}_N^{n+1} + (x\cdot\nabla u_N^n)\overline{u}_N^n]dx$, (3.12) is equivalent to

$$\begin{aligned}
\frac{S_n - S_{n-1}}{\Delta t} &= -2dIm\int_\Omega \overline{\varphi}^n\frac{u_N^{n+1}-u_N^n}{\Delta t}dx - 2dIm\int_\Omega \overline{\varphi}^{n-1}\frac{u_N^{n}-u_N^{n-1}}{\Delta t}dx \\
&\quad -4Im\int_\Omega (x\cdot\nabla\overline{\varphi}^n)\frac{u_N^{n+1}-u_N^n}{\Delta t}dx \\
&\quad -4Im\int_\Omega (x\cdot\nabla\overline{\varphi}^{n-1})\frac{u_N^{n}-u_N^{n-1}}{\Delta t}dx \\
&= \Theta_1 + \Theta_2 + \Theta_3 + \Theta_4. \qquad (3.13)
\end{aligned}$$

Because of

$$\begin{aligned}
\Theta_1 &= -2dIm\int_\Omega \overline{\varphi}^n\frac{u_N^{n+1}-u_N^n}{\Delta t}dx \\
&= 2dRe\int_\Omega \overline{\varphi}^n[-\Delta\varphi^n + v(x)\varphi^n - \lambda(\phi(u_N^{n+1},u_N^n)\varphi^n]dx \\
&= 2d\|\nabla\varphi^n\|^2 + 2d\int_\Omega v(x)|\varphi^n|^2dx - 2d\lambda\int_\Omega \phi(u_N^{n+1},u_N^n)|\varphi^n|^2dx, \qquad (3.14)
\end{aligned}$$

similarly

$$\Theta_2 = 2d\|\nabla\varphi^{n-1}\|^2 + 2d\int_\Omega v(x)|\varphi^{n-1}|^2dx - 2d\lambda\int_\Omega \phi(u_N^{n},u_N^{n-1})|\varphi^{n-1}|^2dx, \qquad (3.15)$$

$$\begin{aligned}
\Theta_3 &= -4Im\int_\Omega (x\cdot\nabla\overline{\varphi}^n)\frac{u_N^{n+1}-u_N^n}{\Delta t}dx \\
&= 4Re\int_\Omega (x\cdot\nabla\overline{\varphi}^n)[-\Delta\varphi^n + v(x)\varphi^n - \lambda(\phi(u_N^{n+1},u_N^n)\varphi^n]dx
\end{aligned}$$

$$
\begin{aligned}
&\le (4-2d)\|\nabla\varphi^n\|^2 - 2d\int_\Omega v(x)|\varphi^n|^2dx - 2\int_\Omega (x\cdot\nabla v(x))|\varphi^n|^2dx \\
&\quad -2\lambda\int_\Omega (x\cdot\nabla|\varphi^n|^2)\phi(u_N^{n+1},u_N^n)dx. \qquad (3.16)
\end{aligned}
$$

Similar to the estimation of Θ_3, we have

$$
\begin{aligned}
\Theta_4 &\le (4-2d)\|\nabla\varphi^{n-1}\|^2 - 2d\int_\Omega v(x)|\varphi^{n-1}|^2dx - 2\int_\Omega (x\cdot\nabla v(x))|\varphi^{n-1}|^2dx \\
&\quad -2\lambda\int_\Omega (x\cdot\nabla|\varphi^{n-1}|^2)\phi(u_N^{n},u_N^{n-1})dx. \qquad (3.17)
\end{aligned}
$$

Consequently, we obtain the estimation (3.13) in terms of (3.14),(3.15) and (3.16),

$$
\begin{aligned}
\frac{S_{n+1}-S_n}{\Delta t} \le\ & 4(\|\nabla\varphi^{n-1}\|^2+\|\nabla\varphi^n\|^2) - 2\int_\Omega (x\cdot\nabla v(x))(|\varphi^{n-1}|^2+|\varphi^n|^2)dx \\
& -2d\lambda\int_\Omega \phi(u_N^{n+1},u_N^n)|\varphi^n|^2dx - 2d\lambda\int_\Omega (x\cdot\nabla|\varphi^n|^2)\phi(u_N^{n+1},u_N^n)dx \\
& -2d\lambda\int_\Omega \phi(u_N^{n},u_N^{n-1})|\varphi^{n-1}|^2dx \\
& -2\lambda\int_\Omega (x\cdot\nabla|\varphi^{n-1}|^2)\phi(u_N^{n+1},u_N^n)dx. \qquad (3.18)
\end{aligned}
$$

In the following, we concern about the nonlinear term above, that is

$$
\begin{aligned}
& -2d\lambda\int_\Omega \phi(u_N^{n+1},u_N^n)|\varphi^n|^2dx - 2\lambda\int_\Omega (x\cdot\nabla|\varphi^n|^2)\phi(u_N^{n+1},u_N^n)dx \\
&= 2\lambda\int_\Omega (x\cdot\nabla(\phi(u_N^{n+1},u_N^n)|\varphi^n|^2)dx - 2\lambda\int_\Omega (x\cdot\nabla|\varphi^n|^2)\phi(u_N^{n+1},u_N^n)dx \\
&= 2\lambda\int_\Omega (x\cdot\nabla\phi(u_N^{n+1},u_N^n))|\varphi^n|^2dx \\
&= \frac{\lambda}{p+1}\int_\Omega (x\cdot\nabla(|u_N^{n+1}|^{2p}+|u_N^{n+1}|^{2p-2}|u_N^n|^2+\cdots+|u_N^n|^{2p})\frac{1}{2}|u_N^{n+1}+u_N^n|^2dx \\
&\le \frac{\lambda}{p+1}\int_\Omega (x\cdot\nabla(|u_N^{n+1}|^{2p}+|u_N^{n+1}|^{2p-2}|u_N^n|^2+\cdots+|u_N^n|^{2p})(|u_N^{n+1}|^2+|u_N^n|^2)dx,
\end{aligned}
$$

by Young's inequality,

$$
\begin{aligned}
& |u_N^{n+1}|^{2p-2k}|u_N^n|^{2k}(|u_N^{n+1}|^2+|u_N^n|^2) \\
&= |u_N^{n+1}|^{2p-2k+2}|u_N^n|^{2k} + |u_N^{n+1}|^{2p-2k}|u_N^n|^{2k+2} \\
&\le \frac{p-k+1}{p+1}|u_N^{n+1}|^{2p+2} + \frac{k}{p+1}|u_N^n|^{2p+2} + \frac{p-k}{p+1}|u_N^{n+1}|^{2p+2} + \frac{k+1}{p+1}|u_N^n|^{2p+2} \\
&= \frac{2p-2k+1}{p+1}|u_N^{n+1}|^{2p+2} + \frac{2k+1}{p+1}|u_N^n|^{2p+2} (k=0,1,\cdots,p).
\end{aligned}
$$

Consequently,

$$2\lambda \int_{\Omega} (x \cdot \nabla\phi(u_N^{n+1}, u_N^n))|\varphi^n|^2 dx \leq -\frac{pd\lambda}{p+1} \int_{\Omega} (|u_N^{n+1}|^{2p+2} + u_N^n|^{2p+2}) dx. \tag{3.19}$$

Moreover, due to the conservation laws $E_1(u_N^n) = E_1(u_N^0)$ and their expressions, together with the hypothesis in theorem 3.2, (3.18) can be rewritten as

$$\begin{aligned} \frac{S_{n+1} - S_n}{\Delta t} \leq\; & 2(\|u_N^{n+1}\|^2 + 2\|u_N^n\|^2 + \|u_N^{n-1}\|^2) \\ & - \int_{\Omega} (x \cdot \nabla v(x))(|u_N^{n+1}|^2 + 2|u_N^n|^2 + |u_N^{n-1}|^2) dx \\ & - \frac{pd\lambda}{p+1} (\|u_N^{n+1}\|_{L^{2p+2}}^{2p+2} + 2\|u_N^n\|_{L^{2p+2}}^{2p+2} + \|u_N^{n-1}\|_{L^{2p+2}}^{2p+2})) \\ \leq\; & 16E_1(u_N^0) - \int_{\Omega} (2v(x) + x \cdot \nabla v(x))(|u_N^{n+1}|^2 + 2|u_N^n|^2 + |u_N^{n-1}|^2) dx \\ & -\lambda \frac{dp-4}{p+1} (\|u_N^{n+1}\|_{L^{2p+2}}^{2p+2} + 2\|u_N^n\|_{L^{2p+2}}^{2p+2} + \|u_N^{n-1}\|_{L^{2p+2}}^{2p+2}) \\ \leq\; & 16E_1(u_N^0). \end{aligned} \tag{3.20}$$

Consequently,

$$S_n \leq S_0 + 16E_1(u_N^0)T. \tag{3.21}$$

Moreover, because that an equivalent expression of (3.20) can be expressed as

$$\frac{W_{n+1} - 2W_n + W_{n-1}}{\Delta t^2} = \frac{S_{n+1} - S_n}{\Delta t} \leq 16E_1(u_N^0), \tag{3.22}$$

$$W_n \leq W_0 + S_0 T + 8E_1(u_N^0)T^2. \tag{3.23}$$

If $\lim_{t_n \to T^*} W_n = \lim_{t_n \to T^*} \int_{\Omega} |x|^2 |u_N^n|^2 dx = 0$, the right-hand side of the last inequality becomes negative for $T > T^*$ provided one of assumptions (1)-(3) of Theorem 3.2 holds.

Theorem 3.3. *Assume that the conditions of Theorem 3.2 hold. If*

$$\lim_{t_n \to T^*} \int_{\Omega} |x|^2 |u_N^n|^2 dx = 0, \tag{3.24}$$

where T^* is the smallest positive zero of W_n, then

$$\lim_{t_n \to T^*} \|u_N^n\|_{L^q} = 0 \qquad \text{if} \quad 1 \leq q < 2, \tag{3.25}$$

and

$$\lim_{t_n \to T^*} \|u_N^n\|_{L^q} = +\infty \qquad \text{if} \quad 2 < q \leq +\infty. \tag{3.26}$$

Proof. Let $q \in [1, 2)$ be a fixed number, choose a constant α such that

$$0 < \alpha < min\{q, \frac{d}{2}(2-q)\},$$

then apply the Hölder inequality two times, we get with regard to (3.2)

$$\begin{aligned}
\int_\Omega |u_N^n|^q dx &= \int_\Omega |x|^{-\alpha}|x|^{\alpha}|u_N^n|^q dx \\
&\le (\int_\Omega |x|^{-\frac{2\alpha}{2-q}} dx)^{(2-q)/2}(\int_\Omega |x|^{2\alpha/q}|u_N^n|^2 dx)^{q/2} \\
&= c(\int_\Omega |x|^{2\alpha/q}|u_N^n|^{2\alpha/q}|u_N^n|^{2(1-\alpha/q)} dx)^{q/2} \\
&\le c(\int_\Omega |x|^2|u_N^n|^2 dx)^{\alpha/2}(\int_\Omega |u_N^n|^2 dx)^{(q-\alpha)/2} \\
&= c(\int_\Omega |x|^2|u_N^n|^2 dx)^{\alpha/2}\|u_N^0\|^{q-\alpha} dx \to 0,
\end{aligned}$$

as $t_n \to T^*, t_n < T^*$, where c is a positive constant, therefor

$$\lim_{t_n \to T^*} \|u_N^n\|_{L^q} = 0 \qquad \text{if } 1 \le q < 2.$$

If $q > 2$, then we use the Hölder inequality once more to obtain

$$0 < \|u_N^0\|^2 = \|u_N^n\|^2 \le \|u_N^n\|_{L^q}\|u_N^n\|_{L^s},$$

where $s \in [1, 2)$ and $1/q + 1/s = 1$. Noting (3.24) and the assumption that

$$\int_\Omega |x|^2|u_N^n|^2 dx \to 0 \qquad \text{as } \ t_n \to T^*,$$

we conclude

$$\lim_{t_n \to T^*} \|u_N^n\|_{L^q} = +\infty \qquad \text{if } \ 2 < q \le +\infty.$$

Corollary 3.1. *Assume that the conditions of Theorem 3.2 hold. The inequality*

$$\|\nabla u_N^n\| \ge \frac{d}{2}\frac{\|u_N^n\|^2}{\|xu_N^n\|}$$

implies that

$$\lim_{t_n \to T^*} \|\nabla u_N^n\| = +\infty, \tag{3.27}$$

if $\int_\Omega |x|^2|u_N^n|^2 dx \to 0$ *as* $t_n \to T^*$.

Corollary 3.2. *Suppose that* $u_0(x)$ *is radially symmetric and it satisfies the conditions of the theorem 3.2 if* $d \ge 2$. *Then the solutions of (2,1) and (2.2),* *i.e.,* $u_N^n(\forall n = 0, 1, \cdots, M)$*, will blow up in limit time, and the origin is the point of blow-up. That is,* $\forall R > 0$*, it holds*

(i) $\lim_{t_n \to T^*} \|u_N^n\|_{L^\infty(|x|<R)} = +\infty$,

(ii) $\lim_{t_n \to T^*} \|u_N^n\|_{L^q(|x|<R)} = +\infty \quad (2 < q \le \infty)$.

In fact, according to

$$\int_{|x|>R} |u_N^n|^2 dx \le \frac{1}{R} \int_\Omega |x|^2 |u_N^n|^2 dx \to 0 \quad (t_n \to T^*),$$

and

$$\|u_N^n\|_{L^q(|x|<R)}^q \le \|u_N^n\|_{L^2} \|u_N^n\|_{L^\infty(|x|<R)}^{q-2},$$

both (i)and (ii) can be obtained easily.

4 Conclusion

In all, we study the blow-up for initial and Dirichlet boundary-value problem of a class of nonlinear Schrödinger equations with potential. We construct fully discrete spectral method by difference in time and by the spectral scheme in space. We prove that the scheme satisfies the conservation of energy. Then both sufficient conditions about the blow-up approximation and the maximum time interval of the solution existence are obtained. It is proved that the blow-up point in space is the origin for radial symmetry if the solution blows up.

Acknowledgements. The research is supported by Fujian National Science Foundation of China Grant 2008J0198.

References

1. Kelley, P.L.: Self focusing of optical beams. Phy. Rev. Lett. 15, 1005–1008 (1965)
2. Zakharov, V.E.: Collapse of Langmuir waves. Sov. Phys. JETP 35, 909–912 (1972)
3. Bradley, C.C., Sackett, C.A., Tollett, J.J., et al.: Evideence of Bose-Einstein condensation in an atomic gas with attractive interactions. Phys. Rev. Lett. 75, 1687–1690 (1995)
4. Bradley, C.C., Sackett, C.A., Hulet, R.G.: Bose-Einstein condensation of lithium observation of limited condensate number. Phys. Rev. Lett. 78, 985–989 (1997)
5. Carles, R.: Nnolinear Schrödinger equations with repulsive harmonic potential and applications. SIAM J. Math. Anal. 35, 823–843 (2003)
6. Carles, R.: Global existence for nonlinear Shcrödinger equations with quadratic potential. Discrete Contin. Dyn. Syst. 13, 385–398 (2005)

7. Beresytcki, H., Cazenave, T.: Instabilité des états stationnaires dans les équations de Schrödinger de Klein-Gordon non linéairees. C. R. Acad. Sci. Paris 293, 489–492 (1981)
8. Sulem, C., Sulem, P.L.: The nonlinear Schrödinger equation self-focusing and wave collapse. Applied Mathematical. Springer, New York (1999)
9. Taha, T.R., Ablowitz, M.J.: Analytical and numerical aspects of certain nonlinear evolution equations II. Numerical. Nonlinear Schrödinger Equation. Jour. Comp. Phys. 55, 203–230 (1984)
10. Delfour, M., Fortin, M., Payre, G.: Finite difference solutions of a Nonlinear Schrödinger Equation. J. Comp. Phys. 44, 277–288 (1981)
11. Tourigny, Y., Morris, J.: An investigation into the effect of product approximation in the numerical solution of the cubic Nonlinear Schrödinger Equation. J. Comp. Phys. 76, 103–130 (1988)
12. Weideman, J.A.C., Herbst, B.M.: Split-step methods for the solution of the Nonlinear Schrödinger Equation. SIAM J. Numer. Anal. 23, 485–507 (1986)
13. Akrivis, G.D., Dougalis, V.A., Karakashian, O.A., et al.: Numerical approximation of blow-up of radially symmetric solutions of the nonlinear Schrödinger equation. SIAM J. Sci. Comput. 25, 186–212 (2003)
14. Jiménez, S., Llorente, I.M., Mancho, A.M., et al.: A numerical scheme for the simulation of blow-up in the nonlinear Schrödinger equation. Computing 77(4), 335–346 (2006)

P-Sets and $\bar{F}$ -Data Selection-Discovery

Ming Xiu[1,2], Kai-quan Shi[2], and Li Zhang[2]

College of Engineering Technology,
Jimei University, Xiamen 361021, China
e-mail: jdxml@tom.com
Department of Mathematics and System Sciences,
Shandong University, Jinan 250100, China

Abstract. P-sets (packet sets) are set pair which are composed of internal P-sets X^F (internal packet set X^F) and outer P-sets X^F (outer packet sets X^F), or (X^F ,X^F) is P-sets. P-sets have dynamic characteristics. P-sets which come from finite general sets have new mathematic structure and the concept of P-sets is new. Based on P-sets, the concept of $\bar{F}$ -data, $\bar{F}$ -data selection guideline , $\bar{F}$ -data selection theorem are given. By using these results, applications of $\bar{F}$ -data selection are proposed. P-sets are new method to study data system.

Keywords: P-sets, $\bar{F}$ -data, data selection, data circle theorem, data selection theorem, applications.

1 Introduction

There is an example as this: x_1, x_2, x_3, x_4, x_5 are five apples which have $\alpha_1 = \text{red}, \alpha_2 = \text{sweet}$, then there is a set of apple $X = \{x_1, x_2, x_3, x_4, x_5\}$; in other words, every $x_k \in X$ in X has α_1 and α_2 , $k = 1, 2, \cdots, 5$. If α_1 and α_2 are defined as attributes $x_k \in X$, $X = \{x_1, x_2, x_3, x_4, x_5\}$ has attribute set $\alpha = \{\alpha_1, \alpha_2\}$. Obviously, set X and set α are corresponding, or, given set $X = \{x_1, x_2, x_3, x_4, x_5\}$, there must be set $\alpha = \{\alpha_1, \alpha_2\}$; $\forall x_i \in X, x_i$ has $\alpha = \{\alpha_1, \alpha_2\}, i = 1, 2, \cdots, 5$. If attribute $\alpha_3 = 200\text{g}$ is added to attribute set α , or, $\alpha = \{\alpha_1, \alpha_2\}$ changes into $\alpha^F = \alpha \bigcup \{\alpha_3\} = \{\alpha_1, \alpha_2, \alpha_3\}$, then attribute set $X = \{x_1, x_2, x_3, x_4, x_5\}$ changes into attribute set $X^{\bar{F}} = \{x_1, x_4\}; x_1, x_4 \in X^{\bar{F}}$ have attribute set $\alpha^{\bar{F}}$. There easily to see: If attribute set α is supplemented by attribute set α_3 ,then the set of apple becomes smaller, X changes into $X^{\bar{F}}$, or $X^{\bar{F}} = \{x_1, x_4\} \subseteq \{x_1, x_2, x_3, x_4, x_5\} = X$. This is a simple fact. Further discussion we have: given set $X = \{x_1, x_2, \cdots, x_m\}$, $\alpha = \{\alpha_1, \alpha_2, \cdots, \alpha_k\}$ is attribute set of X . If X is defined as data X , then data X has attribute set α . Adding attributes $\alpha_{k+1}, \alpha_{k+2}, \cdots, \alpha_{k+r}$ to α , we can select data $X^{\bar{F}} = \{x_1, x_2, \cdots, x_n\}$ from X ,we want to get data $X^{\bar{F}}$, where $X^{\bar{F}} = \{x_1, x_2, \cdots, x_n\}$ $\subseteq \{x_1, x_2, \cdots, x_m\} = X$, $\alpha^{\bar{F}} = \alpha \bigcup \{\alpha_{k+1}, \alpha_{k+2}, \cdots, \alpha_{k+r}\} = \{\alpha_1, \alpha_2, \cdots, \alpha_k, \alpha_{k+1}, \cdots, \alpha_{k+r}\}$ is attribute set of data $X^{\bar{F}}$. This example is similar to P-sets ($X^{\bar{F}}$, X^F)[1,2] . In 2008, Ref. [1,2] introduced dynamic characteristic to finite general set X (cantor set X), the finite general set X is improved and P-sets (packet set) are proposed, the

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 791–799.
springerlink.com

structure and applications of P-sets are given. P-sets are new module to study data system and information system.

Using structure of P-sets, the paper gives concept $\overline{F}$ -data, $\overline{F}$ -data selection guideline and $\overline{F}$ -data selection theorem. The results given in this paper can be used in many applied fields of data system and information system.

For the convenience and keeping the contents integral, the structure of P-sets is introduced to section 2 as the theory basis of the discussion of this paper. The more concepts and applications of P-sets, please to see Ref. [1-5].

2 The Structure and Characteristic of P-sets

Assumption 1. *X is a finite general set on U, U is a finite element universe, V is a finite attribute universe.*

Given a general set $X = \{x_1, x_2, \cdots, x_m\} \subset U$, and $\alpha = \{\alpha_1, \alpha_2, \cdots, \alpha_k\} \subset V$ is attribute set of X, $X^{\overline{F}}$ is called internal packet sets of X, called internal P-sets for short, moreover

$$X^{\overline{F}} = X - X^{-}. \tag{1}$$

X^{-} is called $\overline{F}$ -element deleted set of X, moreover

$$X^{-} = \{x \mid x \in X, \overline{f}(x) = u \overline{\in} X, \overline{f} \in \overline{F}\}, \tag{2}$$

if the attribute set α^F of $X^{\overline{F}}$ satisfies

$$\alpha^F = \alpha \cup \{\alpha' \mid f(\beta) = \alpha' \in \alpha, f \in F\}, \tag{3}$$

where $X^{\overline{F}} \neq \phi, \beta \in V, \beta \overline{\in} \alpha, f \in F$ turns β into $f(\beta) = \alpha' \in \alpha$.

Given a general set $X = \{x_1, x_2, \cdots, x_m\} \subset U$, and $\alpha = \{\alpha_1, \alpha_2, \cdots, \alpha_k\} \subset V$ is attribute set of X, X^F is called outer packet sets of X, called outer P-sets for short, moreover

$$X^F = X \cup X^{+}. \tag{4}$$

X^{+} is called F -element supplemented set, moreover

$$X^{+} = \{u \mid u \in U, u \overline{\in} X, f(u) \in x' \in X, f \in F\}, \tag{5}$$

if the attribute set $\alpha^{\overline{F}}$ of X^F satisfies

$$\alpha^{\overline{F}} = \alpha - \{\beta_i \mid \overline{f}(\alpha_i) = \beta_i \overline{\in} \alpha, \overline{f} \in \overline{F}\}, \tag{6}$$

where $\alpha^{\overline{F}} \neq \phi, \alpha_i \in \alpha, \overline{f} \in \overline{F}$ turns α_i into $\overline{f}(\alpha_i) = \beta_i \overline{\in} \alpha$.

The set pair which are composed of internal P-sets $X^{\overline{F}}$ and outer P-sets X^{F} are called P-sets (packet sets) generated by general set X , called P-sets for short, moreover

$$(X^{\overline{F}}, X^{F}), \tag{7}$$

where: General set X is ground set of $(X^{\overline{F}}, X^{F})$.

As *P*-sets have dynamic characteristic, the general representation of *P*-sets is:

$$\{(X_i^{\overline{F}}, X_j^{F}) \mid i \in \mathrm{I}, j \in \mathrm{J}\}, \tag{8}$$

where : I, J are index sets, formula (8) is the representation of set pair family of P-sets.

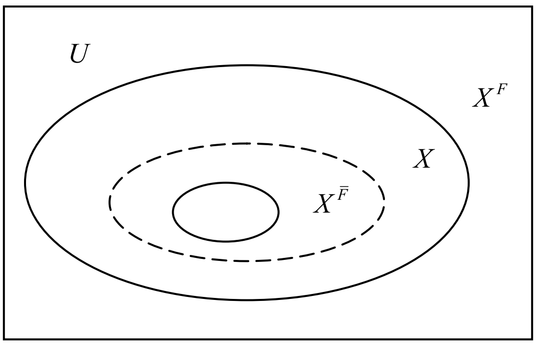


Fig. 1. Gives intuitive graphical representation of *P*-sets. $X \subset U$ is a finite general set on $X = \{x_1, x_2, \cdots, x_q\}$, $X^{\overline{F}}$ is internal *P*-sets of $X \subset U, X^{\overline{F}} = \{x_1, x_2, \cdots, x_p\}, p \le q$, X^F is outer *P*-sets of $X \subset U, X^F = \{x_1, x_2, \cdots, x_r\}$, $q \le r$, $(X^{\overline{F}}, X^F)$ is *P*-sets. $X^{\overline{F}}$ is expressed in thick line, X is expressed in thin line, X^F is expressed in dashed line.

Theorem 1. *If* $\overline{F} = F = \phi$*, then P-sets* $(X^{\overline{F}}, X^F)$ *and general set* X *satisfy*

$$(X^{\overline{F}}, X^F)_{\overline{F}=F=\varnothing} = X . \tag{9}$$

Proof. If $\overline{F} = \phi$,then formula (2) changes into $X^- = \{x \mid x \in X, \overline{f}(x) = u \,\overline{\in}\, X,\ \overline{f} \in \overline{F}\} = \phi$, formula (1) changes into $X^{\overline{F}} = X - X^- = X$, If $F = \phi$, then formula (5) changes into $X^+ = \{u \mid u \in U, u \,\overline{\in}\, X, f(u) = x' \in X, f \in F\} = \phi$, formula (4) changes into $X^F = X \cup X^+ = X$. *P*-sets $(X^{\overline{F}}, X^F)$ changes into X , then there is formula (10).

Formula (10) proposes that under the condition of $\overline{F} = F = \phi$, *P*-sets $(X^{\overline{F}}, X^{F})$ turns back to "origin" of general set , in other words, *P*-sets have lost dynamic characteristics, actually, *P*-sets $(X^{\overline{F}}, X^{F})$ is general set X .

Theorem 2. *If* $\overline{F} = F = \phi$*, then the set pair family of P-sets* $\{(X_i^{\overline{F}}, X_j^{F}) \mid i \in \mathrm{I}, j \in \mathrm{J}\}$ *and general set* X *satisfy:*

$$\{(X_i^{\overline{F}}, X_j^{F}) \mid i \in \mathrm{I}, j \in \mathrm{J}\}_{\overline{F}=F=\varnothing} = X \ . \tag{10}$$

Formula (11) proposes that under the condition of $\overline{F} = F = \phi$, every $X_i^{\overline{F}}$, X_j^{F} turns back to "origin" of general set , or $\{(X_i^{\overline{F}}, X_j^{F}) \mid i \in \mathrm{I}, j \in \mathrm{J}\}$ turns back to "origin" of general set . The proof of theorem 2 is similar to theorem 1's, so the proof is omitted.

It is should be pointed out here:

1°. The characteristic of formulas (1), (3), (4), (6) is similar to the structure of $T = T + 1$ in computer memory. In the field of computer science, $T = T + 1$ is a simple, common concept, people are familiar with it. They have dynamic characteristic. The dynamic characteristic of formula (3) is that β_1 changes into $f(\beta_1) = \alpha_1'$, α_1' comes into α and α_1^F is obtained, or, $\alpha_1^F = \alpha \cup \{f(\beta_1)\} = \alpha \cup \{\alpha_1'\} = \{\alpha, \alpha_1'\}$. If $\alpha = \alpha_1^F$, β_2, β_3 change into $f(\beta_2) = \alpha_2', f(\beta_3) = \alpha_3'$; α_2', α_3' come into α_1^F and α_2^F is obtained, or, $\alpha_2^F = \alpha \cup \{\alpha_2', \alpha_3'\} = \alpha_1^F \cup \{\alpha_2', \alpha_3'\} = \{\alpha, \alpha_1'\} \cup \{\alpha_2', \alpha_3'\} = \{\alpha, \alpha_1', \alpha_2', \alpha_3'\}$. If $\alpha = \alpha_2^F$, β_4 changes into $f(\beta_4) = \alpha_4'$, α_4' comes into α_2^F , or, $\alpha_3^F = \alpha \cup \{\alpha_4'\} = \alpha_2^F \cup \{\alpha_4'\} = (\{\alpha, \alpha_1'\} \cup \{\alpha_2', \alpha_3'\}) \cup \{\alpha_4'\} = \{\alpha, \alpha_1', \alpha_2', \alpha_3', \alpha_4'\}$, and so on, apparently, formula (3) is similar to the structure of $T = T + 1$. It is wrong to use the viewpoint of "static" or "classical mathematics" to understand the four formulas.

2°. $F = \{f_1, f_2, \cdots, f_m\}$, $\overline{F} = \{\overline{f}_1, \overline{f}_2, \cdots, \overline{f}_n\}$ are element transfer families [1-3], $f \in F, \overline{f} \in \overline{F}$ are element transfers[1-3], The characteristic of $f \in F$ is: $u \in U, u \overline{\in} X, f \in F$ changes u into $f(u) = x' \in X$,or $\beta \in V, \beta \overline{\in} \alpha, f \in F$ changes β into $f(\beta) = \alpha' \in \alpha$. The characteristic of $\overline{f} \in \overline{F}$ is: $x \in X, \overline{f} \in \overline{F}$ changes x into $\overline{f}(x) = u \overline{\in} X$, or $\alpha_i \in \alpha$, $\overline{f} \in \overline{F}$ changes α_i into $\overline{f}(\alpha_i) = \beta_i \overline{\in} \alpha$. Obviously, $f \in F, \overline{f} \in \overline{F}$ are given functions (function is a transformation or mapping).

3°. In formula (3): $\alpha^F = \alpha \cup \{\alpha' \mid f(\beta) = \alpha' \in \alpha, f \in F\}$, $\{\alpha' \mid f(\beta) = \alpha' \in \alpha, f \in F\}$ is composed of attributes which are added to α , $\{\alpha' \mid f(\beta) = \alpha' \in \alpha,$

$f \in F\}$ and α which hasn't been supplemented by attributes satisfy that $\{\alpha' \mid f(\beta) = \alpha' \in \alpha, f \in F\} \bigcap \alpha = \phi$. For example: $\alpha = \{\alpha_1, \alpha_2, \alpha_3\}$, $\{\alpha' \mid f(\beta) = \alpha' \in \alpha, f \in F\} = \{\alpha'_1, \alpha'_2, \alpha'_3\}, \alpha_k \neq \alpha'_k, k = 1,2,3$; apparently,as α and $\{\alpha' \mid f(\beta) = \alpha' \in \alpha, f \in F\}$ have no common elements, $\alpha \bigcap \{\alpha' \mid f(\beta) = \alpha' \in \alpha, f \in F\} = \phi$. So the conclusion of "As $\{\alpha' \mid f(\beta) = \alpha' \in \alpha, f \in F\} \subseteq \alpha$, $\alpha^F = \alpha \bigcup \{\alpha' \mid f(\beta) = \alpha' \in \alpha, f \in F\} = \alpha$" is wrong.

4°. Formulas (1)-(3) give concepts as this: Some attributes are deleted from X, X generates internal P-sets $X^{\bar{F}}$, it is equivalent to attribute set α of X being supplemented by new attributes, α generates α^F, $\alpha \subseteq \alpha^F$ .Or, if α_1^F, α_2^F are attribute sets of $X_1^{\bar{F}}, X_2^{\bar{F}}$ respectively, moreover $\alpha_1^F \subseteq \alpha_2^F$, then $X_2^{\bar{F}} \subseteq X_1^{\bar{F}}$. $\{\alpha' \mid f(\beta) = \alpha' \in \alpha, \ f \in F\}$ in formula (3) isn't attribute set of X^- which is composed of elements deleted from X, or $\{\alpha' \mid f(\beta) = \alpha' \in \alpha, f \in F\}$ is attribute set of X^-, X^- is formula (2).

The reason and evidence for the existence of set pair family $\{(X_i^{\bar{F}}, X_j^F) \mid i \in \mathrm{I}, j \in \mathrm{J}\}$

Given finite general set $X = \{x_1, x_2, \cdots, x_m\} \subset U$, $\alpha = \{\alpha_1, \alpha_2, \cdots, \alpha_k\} \subset V$ is attribute set of X. If α is supplemented by some attributes at the same time some attributes are deleted from it, then α changes into $\alpha_1^F, \alpha_1^{\bar{F}}$, $\alpha_1^F \neq \alpha_1^{\bar{F}}$, $\alpha \subseteq \alpha_1^F, \alpha_1^{\bar{F}} \subseteq \alpha$. By formulas (1)-(7), there is P-sets($X_1^{\bar{F}}, X_1^F$) of X .If this process continues, then α changes into $\alpha_2^F, \alpha_2^{\bar{F}}, \alpha_2^F \neq \alpha_2^{\bar{F}}, \alpha \subseteq \alpha_2^F, \alpha_2^{\bar{F}} \subseteq \alpha$ respectively. By formulas (1)-(7) , there is P-sets($X_2^{\bar{F}}, X_2^F$) of X ,and so on. These a string of set pairs($X_i^{\bar{F}}, X_i^F$) compose set pair family of formula (8).

By the concepts in section 2, section 3 gives:

3 $\bar{F}$-Data and $\bar{F}$-Data Selection Theorem

Assumption. *$X, X^{\bar{F}}, X^F$ in section **2** are expressed by $(x), (x)^{\bar{F}}, (x)^F$ respectively, or, $(x) = X, (x)^{\bar{F}} = X^{\bar{F}}$, $(x)^F = X^F$, in order not to cause confusion and misunderstanding.*

Definition 1. *(x) is called data on U ,moreover*

$$(x) = \{x_1, x_2, \cdots, x_q\}, \tag{11}$$

$\forall\, x_i \in (x)$ is called data element of (x), $i = 1, 2, \cdots, q$.

If (x) *has attribute set* α *, moreover*

$$\alpha = \{\alpha_1, \alpha_2, \cdots, \alpha_k\}. \tag{12}$$

Definition 2. $(x)^{\bar{F}}$ *is called* $\bar{F}$ *-data generated by* (x)*,moreover*

$$(x)^{\bar{F}} = \{x_1, x_2, \cdots, x_p\}, \tag{13}$$

$\forall x_j \in (x)^{\bar{F}}$ *is called data element of* $(x)^{\bar{F}}$.

If the attribute set α^F *of* $(x)^{\bar{F}}$ *and attribute set* α *of* (x) *satisfy*

$$\alpha^F = \alpha \cup \{\alpha' \mid f(\beta) = \alpha' \in \alpha, f \in F\}. \tag{14}$$

Definition 3. y *is called characteristic value set of* (x)*, moreover*

$$y = \{y_1, y_2, \cdots y_q\}. \tag{15}$$

$y^{\bar{F}}$ *is called characteristic value set of* $\bar{F}$ *-data* $(x)^{\bar{F}}$*, moreover*

$$y^{\bar{F}} = \{y_1, y_2, \cdots, y_p\}, \tag{16}$$

where: $\forall y_i \in y$ *is characteristic value of data element* $x_i \in (x)$, $\forall y_j \in y^{\bar{F}}$ *is characteristic value of data element* $x_j \in (x)^{\bar{F}}$, $y_i, y_j \in \mathrm{R}$, R *is real number set.*

Definition 4. $\bigcirc$ *is called data circle generated by data* (x) *which considers coordinate origin O as the center and considers* ρ *as the radius , if*

$$\rho = \| y \| / \| y \|, \tag{17}$$

where: $\| y \| = (y_1^2 + y_2^2 + \cdots + y_q^2)^{1/2}$ *is 2-Norm of vector* $y = (y_1, y_2, \cdots, y_q)^T$ *generated by characteristic value set* $y = \{y_1, y_2, \cdots, y_q\}$, $y = \{y_1, y_2, \cdots, y_q\}$ *is generated by formula (15).*

Definition 5. $\bigcirc^{\bar{F}}$ *is called* $\bar{F}$ *-data circle generated by* $\bar{F}$ *-data* $(x)^{\bar{F}}$ *which considers coordinate origin O as the center and considers* $\rho^{\bar{F}}$ *as the radius ,if*

$$\rho^{\bar{F}} = \| y^{\bar{F}} \| / \| y \|, \tag{18}$$

where : $\| y^{\bar{F}} \| = (y_1^2 + y_2^2 + \cdots + y_p^2)^{1/2}$ *is 2-Norm of vector* $y^{\bar{F}} = (y_1, y_2, \cdots, y_p)^T$, $y^{\bar{F}}$ *is vector generated by formula (15).*

By definitions 1-6, there is

Theorem 3 (Internal-data circle theorem of $\overline{F}$ -data selection) *If* $(x)^*$ *is* $\overline{F}$ *-data of* (x)*, then data circle* $\bigcirc^*$ *generated by* $(x)^*$ *and data circle* $\bigcirc$ *generated by* (x) *satisfy*

$$\bigcirc^* \subset \bigcirc, \tag{19}$$

Where:"$\subset$"in formula (19) denotes that $\bigcirc^*$ is surrounded.

Corollary 1 *Given data* $(x)', (x)''$*, if data circle* $\bigcirc'$ *generated by* $(x)'$ *is internal circle of data circle* $\bigcirc''$ *generated by* $(x)''$*,or*

$$\bigcirc' \subset \bigcirc'', \tag{20}$$

then $(x)'$ *is* $\overline{F}$ *-data of* $(x)''$, $(x)' = (x)^{\overline{F}}$.

Theorem 4. (Attribute set theorem of $\overline{F}$ -data selection) *Suppose* α^* *is attribute set of data* $(x)^*$*, and* α *is attribute set of data* (x)*, the necessary and sufficient condition of* $(x)^*$ *being* $\overline{F}$ *-data of* (x) *is*

$$\alpha^* - \alpha \neq \phi. \tag{21}$$

Corollary 2. $\overline{F}$ *-data* $(x)^{\overline{F}}$ *and data* (x) *satisfy*

$$\text{card}((x)^{\overline{F}}) - \text{card}\left((x)\right) \leq 0, \tag{22}$$

where card=cardinal number.

By theorems 3, 4, corollaries 1, 2, we can get easily:

$\overline{F}$ -Data Selection Guideline

Adding attribute α_i' to attribute set α, α generates α^F, data $(x)^{\overline{F}}$ which has attribute set α^F is $\overline{F}$ -data of (x).

Selection guideline give us a method to select $\overline{F}$ -data from data (x).

Using concepts in section 2,3, section 4 gives:

4 Application of $\overline{F}$ -Data Selection

Example of this section is from data security transmission-identification system, A is transmission side of data (x), B is recipient side of data (x). Important data

$(x)^*$ hides in data (x), $(x)^* \subset (x)$. α is attribute set of data (x). Data (x) and its attribute set α are included in table 1.

Table 1. Data (x) and its attribute set α

(x)	x_1	x_2	x_3	x_4	x_5	x_6	x_7
α	α_1	α_2	α_3				

where: the names of attributes $\alpha_1, \alpha_2, \alpha_3$ are omitted, x_i is data element of data (x), $(x) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$, $\alpha = \{\alpha_1, \alpha_2, \alpha_3\}$.

Security Transmission Assumption of Data (x)

α_4', α_5' are attributes added to α, or α is supplemented with attributes α_4', α_5', α changes into $\alpha^F = \alpha \cup \{\alpha_4', \alpha_5'\} = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4', \alpha_5'\}$. α_4', α_5' have characteristic values γ_1, γ_2; $\gamma_1, \gamma_2 \in \mathrm{N}^+$. By γ_1, γ_2, A and B select elliptic curve [12] jointly

$$y^2 = x^3 + ax + b \,. \tag{23}$$

So there is the point $(\gamma_1, \gamma_2) \in E(K) = \{(x, y) \mid x, y \in K, \rho(x, y) = 0\} \cup \{O\}$ which is constituted by γ_1, γ_2.

1°. A uses public key $P_B = n_B G$ in encryption of B to change (γ_1, γ_2) into ciphertext $C = (c_1, c_2)$, A transmits C to B.

2°. B accepts C, B uses secret key n_B in encryption to get (γ_1, γ_2) from C.

The algorithm process of 1°, 2°, the secret key n_A of A and the selection of public key $P_A = n_A G$ in encryption; the secret key n_B of B and the selection of public key $P_B = n_B G$ in encryption, please to see Ref. [11,12], these discussions are omitted here.

3°. B selects data $(x)'$ from (x).

B accepts data $(x) = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$ in table 1, attribute set $\alpha = \{\alpha_1, \alpha_2, \alpha_3\}$, B gets attributes α_4', α_5' from $1^\circ - 2^\circ$, B proposes:

$$\begin{aligned} \alpha^F &= \alpha \cup \{\alpha_4', \alpha_5'\} \\ &= \{\alpha_1, \alpha_2, \alpha_3, \alpha_4', \alpha_5'\} \end{aligned}$$

Using α^F , B accepts data $(x)'$ from (x)

$(x)' = \{x_2, x_3, x_7\}$

Where: $(x)' = \{x_2, x_3, x_7\} \subseteq \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\} = (x)$.

5 Discussion

The paper uses the structure and characteristics of P-sets to give the discussions of $\bar{F}$ -data and $\bar{F}$ -data selection, applications on data security transmission are given. P-sets are new mathematical model to study dynamic data system and dynamic information system, P-sets have great use of space, especially in computer science and information science.

Acknowledgements. This work was supported by the Foundation of Educational Department of Fujian Province, China (Grant No. JA09151, JB08177).

References

1. Shi, K.Q.: P-sets and its applications. An International Journal Advances in Systems Science and Application 9(2), 209–219 (2009)
2. Shi, K.Q.: P-sets. Journal of Shandong University (Natural Science) 43(11), 77–84 (2008)
3. Shi, K.Q., Zhang, L.: Internal P-sets and data outer-recovery. Journal of Shandong University (Natural Science) 44(4), 8–14 (2009)
4. Tang, J.H., Chen, B.H., Shi, K.Q.: P-sets and $(\bar{F}, F)$ -data generation-identification. Journal of Shandong University (Natural Science) 44(11), 83–92 (2009)
5. Yu, X.Q.: Identification and filtering of P-sets. Journal of Shandong University (Natural Science) 45(1), 94–98 (2010)
6. Yu, X.Q.: Area characteristic of P-sets. Journal of Shandong University (Natural Science) 45(2), 99–102 (2010)
7. Zhang, F., Chen, P., Zhang, L.: P-separation of P-sets and its applications. Journal of Shandong University (Natural Science) 45(3), 71–75 (2010)
8. Shi, K.Q., Li, X.H.: Camouflaged information identification and its application. An International Journal Advances in System Science and Applications 10(2), 208–216 (2010)
9. Zhang, L., Cui, Y.Q.: Outer P-sets and its P-separation theorems. An International Journal Advances in System Science and Applications 10(2), 229–236 (2010)
10. Lin, H.K., Li, Y.Y.: P-sets and its P-separation theorems. An International Journal Advances in System Science and Applications 10(2), 245–251 (2010)
11. Zhang, G.Y., Li, E.Z.: Information gene and its information knock-out/knock-in. Science in China F: Information Science 51(7), 924–935 (2010)
12. Shi, K.Q., Zhao, J.L.: Function S-rough sets and security-authentication of hiding law. An International Journal Advances in System Science and Applications 10(2), 924–935 (2008)
13. Koblitz, N.: Elliptic curve cryptosystems. Mathematics and Computer (1), 203–209 (1987)

AI-VCR Addition Computing without Overflowing

Qiu-sun Ye

Department of Mathematics and Computer Science Engineering,
Wuyi University Wuyishan
354300, Fujian, P.R. of China
e-mail: qsye2005@yahoo.com.cn

Abstract. In Computation of Numbers, addition is rather simple and universal computation of arithmetic. However, no matter what advanced computer in the world, and no matter what backward calculator(such as abacus in China etc.) in people's hand, to do sum of addition is always a NP problem in all the time. Such as capacities of numbers in variables, precise computing (PC) of which is concerned with a certain algorithm, computing precision (CP) of which is limited to data bits of CPU word size in computer, perturbation motion(PM) of which is owing to some errors in computing of numbers, etc. This paper gives a realization of novel addition with AI-VCR computation in computer, and its reality solutions of avoiding the sum value's overflowing after many times additions, PM in arithmetic computation, insufficient valid figures, etc.

Keywords: AI-VCR (Variable Carrying Rules), Addition, VCO (Variable Capacities of Overflowing), PM (Perturbation Motion), CPU (Central Processing Units).

1 Forewords

Nowadays, in international Mathematics and Computer Science, all researches on Numbers are mechanized [1] and only confined to same varying rules of FCN (Fixed Carrying Numbers). For example, 10-carrying numbers (D, Decimal numbers), it was used early in SHANG dynasty of China (From B.C. century 17 to 11), until A.D. century 6 it had been used in unity in all over the world [2], today it has been mainly used in unity for 15 centuries (more than 1500 years) in the world; And 2-carrying numbers (B, Binary numbers) was used in main and unity in an inner Computer since 1946. Besides, 5-carrying numbers, 6-carrying numbers, 8-carrying numbers (Q, Octonal numbers), 16-carrying numbers (H, Hexadecimal numbers), etc. All FCN varying rules are called FCR(Fixed Carrying Rules), namely, in a FCN, neighbor Figures' computation rule to be a Same: Binary numbers' Same rules are "Plus 2 as 1, Borrow/Lend 1 as 2", 5-carrying numbers' Same rules are "Plus 5 as 1, Borrow/Lend 1 as 5", 6-carrying numbers' Same rules are "Plus 6 as 1, Borrow/Lend 1 as 6", Octonal numbers' Same rules are "Plus 8 as 1, Borrow/Lend 1 as 8", Decimal numbers' Same rules are "Plus 10 as 1, Borrow/Lend 1 as 10", Hexadecimal numbers' Same rules are "Plus 16 as 1, Borrow/Lend 1 as 16", and the like.

However, in Role of Engineering in Human Society[3], some researches on Numbers must be AI(Artificial Intelligence) and Different varying rules of VCN

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 801–810.
springerlink.com

(Variable Carrying Numbers) which was promoted by Qiusun Ye in 1995. Namely, in a VCN, neighbor Figures' computation rule to be a Variable. For example, the Numbers of time: 2006 years 8 months 5 days, 1 year=12 or 13(the lunar calendar of Chinese leap year) months, 1 month= 30 or 31 days(Chinese calendar may be 29 or 30 days, specially, in a normal year without including leap month, the February=28 days), 1 day=24 hours, 1 hour=60 minutes, 1 minute=60 seconds. The different Figures(year, month, day, hour, minute and second) computation rules are variable, it is also called VCR(Variable Carrying Rules). Qiusun Ye thought that, running & changing of matter in the world was absolute, and the matter's stopping & fixing was relative; so the Numerical Rule, a certain conversion regularity of computation we ought to abide by in a process of describing quantity of matters would be changed in the movement. In an astronomical yearbook, if scientists(who worked in China and in the other countries) didn't inlay AI-Properties of VCR among the Figures such as year, month, day, hour, minute and second; then we couldn't accurately describe periodic changing of weather in one year of which including 4 quarters with 24 climates, peasants wouldn't know planting time for a variety of plants in nature. Of course, today under a big shed with plastic, peasants could also plant a variety of vegetables in different quarter's climates, but these vegetables' taste wouldn't be so good for people in the course of nature. Numbers [4,5], the objects of its study, is mainly a quantity describing how much/many of matters. The matters ought to be an identical with broad matters in concept of Philosophy. It may be an objective reality matters in general sense such as fire, water, fish, and so on. And it also may be an abstract matters or an appearance of matters like that: sound, city, flood, and the like. To describe how much/many of the above matters mentioned, we can describe them as follows: 1 fish, 2 tons of water, 3 fire shows, 4 sounds, 5 cities, 6 times flood, and so forth. Of course, all of these numbers are integer type here. Sometimes, in need of a practice computation, numbers are also fraction type such as 0.5(or 1/2, decimal number with limited figures 0 and 5); 0.666…(or 2/3, recurring decimal number with limitless figures 6); 0.75(or 3/4, decimal number with limited figures 0, 7 and 5); $(25.513)_6 \div (5)_6 = (3.324111\ldots)_6$, recurring decimal number with limitless figures 1;π=3.1415926…(not recurring decimal number, $\pi$is the Ratio of circumference of a circle to its diameter), etc.

After the 10th Conference of Chinese Association for Artificial Intelligence (CAAI) in 2003, the World Famous Mathematician, the first Top-Prize (¥5,000,000) Winner of National Science & Technology of China in 2000, the 24th Conference Chairman of World Mathematician in 2002, the third International Yifu Shao Mathematics Science Prize ($1,000,000) Winner in 2006, the Academician of Chinese Academy of Sciences (CAS), the Academician of the third World Academy of Sciences, the Pre-President of Chinese Mathematics Society(CMS), the Honor President of CMS, the Advisory Committee Honor Chairman of CAAI, Mr. Wentsun Wu was sure that, the VCN promoted by Mr. Qiusun Ye was a novel broad concept of numbers, there would be indeed too much potential science value of researches & applications on VCN. The Famous Expert of AI, the Advisory Committee Chairman of CAAI, the Pre-President of CAAI, the Honor President of CAAI, professor Xuyan Tu thought that, to research the VCN & its applications, and create Mathematics Theory & Methods of the VCN, it owns very important meanings of academy and very wide value of applications, it may be used not only in Cyphering Science, Communication and Safety of Information, but also in all kinds of variable constructions, variable parameters, constructing models of complex systems and analysis & synthesis, researching & development of new theory or methods & new technology in AI. Thereupon, the Academician of Chinese Academy of Engineering (CAE), Editor in Chief of the journal of Engineering Sciences, Mr. Xuguang Wang was also interested in the VCN, so the invited paper of title as VCN & Its Role of Engineering in Human Society was to be published in the journal of Engineering Sciences (Vol.6, No.1, Mar., 2008). In this paper, it will introduce a novel practical AI-VCR addition computing without overflowing.

2 Numbers Computation of VCN & FCN

We suppose that, $F_{n-1}F_{n-2}\ldots F_0.F_{-1}F_{-2}\ldots F_{-m}$ is an optional FCN of including n-Figures integer and m-Figures fraction, and then its Real number value of the FCN may be computed in computation formula as follows:

$$R_{FCN} = F_{n-1}F_{n-2}\cdots F_0 \bullet F_{-1}F_{-2}\cdots F_{-m} = \sum_{i=-m}^{n-1} F_i (r+1)^i \text{ ,} \tag{1}$$

$$Max(I_{FCN}) = F_{n-1}F_{n-2}\cdots F_0 = (r+1)^n - 1 = r\sum_{i=0}^{n-1}(r+1)^i \text{ ,} \tag{2}$$

$$CP(R_{FCN}) = (r+1)^{-V_D}, m+n \le V_D \left(\text{Digits of valid numbers}\right), \tag{3}$$

$$Max(R_{FCN}) = F_{n-1}F_{n-2}\cdots F_0 \bullet F_{-1}F_{-2}\cdots F_{-m} = (r+1)^n - (r+1)^{-m}, \tag{4}$$

$i \in N$, $m \in N$, $n \in N$, $N = \{1,2,3,\cdots\}$, $I_{FCN} = F_{n-1}F_{n-2}\cdots F_0$, $r = Max(F_i)$, $R_{n-1} = Max(F_{n-1}), R_{n-2} = Max(F_{n-2}), \cdots, R_0 = Max(F_0)$.

In an optional FCN, all the biggest Figures are equal to each other, namely, $R_{n-1} = R_{n-2} = \cdots = R_0 \equiv R = r$. The Figures' Module(*FM*) are equal to r+1. In a FCN integer of n-Figures, its Numbers' Module(*NM*) is the biggest number plus one, namely,

$$NM = Max(I_{FCN}) + 1 = (r+1)^n . \tag{5}$$

Similarly, we suppose that, $F_{n-1}F_{n-2}\ldots F_0.F_{-1}F_{-2}\ldots F_{-m}$ is an optional VCN of including n-Figures integer and m-Figures fraction, and then its Real number value of the VCN may be computed in computation formula as follows:

$$R_{VCN} = F_{n-1}F_{n-2}\cdots F_0 \bullet F_{-1}F_{-2}\cdots F_{-m} = \\ \sum_{k=1}^{n-1} F_k \prod_{i=0}^{k-1}(r_i+1) + F_0 + \sum_{j=-2}^{-m} F_j \prod_{\ell=-1}^{-j+1}(r_\ell+1)^{-1} + F_{-1}, \tag{6}$$

$$k \in N\,, i \in N\,, m \in N\,, n \in N\,, j \in I\,, \ell \in I\,, I = \{IntegerNumbers\}\,.$$

In an optional VCN integer of n-Figures, not all the biggest Figures are equal to each other, the Figures' Module (FM_i) are separately equal to R_i+1, its Numbers' Module (*NM*) is the biggest number plus one, or Multiplication of Rolling with all FMs, namely,

$$FM_i = R_i + 1\,, i \in \{0,1,2,\cdots,n-1\}\,,\ R_i = Max(F_i)\,, \tag{7}$$

$$NM = Max(I_{VCN}) + 1 = 1\underbrace{00\cdots00}_{n-zeros} = \prod_{i=0}^{n-1} FM_i\,, \tag{8}$$

$$CP(R_{VCN}) = \prod_{i=-V_D}^{-1} (r_i + 1)^{-1}, m + n \le V_D \ (\text{Digits of valid numbers}), \tag{9}$$

$$Max(R_{VCN}) = F_{n-1}F_{n-2}\cdots F_0 \bullet F_{-1}F_{-2}\cdots F_{-m} = \prod_{i=0}^{n-1}(r_i + 1) - \prod_{j=-m}^{-1}(r_j + 1)^{-1}\,. \tag{10}$$

3 AI-VCR, VCO, CN, PM, CP and PC

In a VCN, we know that different figures' computation rules are called VCR, the VCR is relatively fixed in the VCN, but the VCN length of figures after making a certain computation(such as addition, etc.) with another number would be variable, so the VCR is always dynamic in changing value of VCN. When the VCR will be designed in Artificial Intelligence (AI), this computation rule is called AI-VCR [6].

In a computer, we know that capacity of number value is always limited for data word size of CPU(Central Processing Unit), the word size of CPU to be expressed with Bits of binary numbers(binary codes such as 0 and 1), such as 1 Byte=8Bits, 2 Bytes=16Bits, 4 Bytes=32Bits, 8 Bytes=64Bits, etc. When a number value is over more than the biggest number (≥*NM*), the number will be overflowing in computer. However, if AI-VCR is used of designing a dynamic VCN, then NM of the VCN may be getting larger so that it is limitless, it results in without overflowing forever. This is also called Variable Capacities of Overflowing (VCO).

In Number Theory of Mathematics Science, Numbers are processed in many kinds of arithmetic signs such as ＋, －, ×, ÷, etc. We called it Computation of Numbers (CN). In CN of traditional FCN and its extensive VCN, Perturbation Motion (PM) is the main origins of errors resulted in computing of numbers. Precise Computing (PC) is concerned with an algorithm of Mathematics, and Computing Precision (CP) is concerned with valid number of Figures in computer, namely the data word size of CPU (Bits of binary numbers).

4 AI-VCR Addition Computing with VCO

An addition computation SUM with i decimal numbers [R, Real decimal numbers including Integer(Whole number) and Fraction] may be described as follows:

$$S = A_1 + A_2 + \cdots + A_i = \sum_{j=1}^{i} A_j \text{ , } j \in N \text{ , } \quad N = \{1, 2, 3, \cdots\} \text{ ,} \tag{11}$$

$$A_1 = F_{n1-1}F_{n1-2} \cdots F_0 \bullet F_{-1}F_{-2} \cdots F_{-m1} = D_{n1} \bullet D_{m1} \text{ , } n1 \in N \text{ , } m1 \in N \text{ , } A_1 \in R \text{ ,}$$

$$A_2 = F_{n2-1}F_{n2-2} \cdots F_0 \bullet F_{-1}F_{-2} \cdots F_{-m2} = D_{n2} \bullet D_{m2} \text{ , } n2 \in N \text{ , } m2 \in N \text{ , } A_2 \in R \text{ ,}$$

……

$$A_i = F_{ni-1}F_{ni-2} \cdots F_0 \bullet F_{-1}F_{-2} \cdots F_{-mi} = D_{ni} \bullet D_{mi} \text{ , } ni \in N \text{ , } mi \in N \text{ , } A_i \in R \text{ .}$$

V_{D1}=n1+m1, V_{D2}=n2+m2, …, V_{Di}=ni+mi, V_D is Valid number for Figures which is limited on the Data word size (Bits of binary numbers) of CPU in computer.

Owing to the Properties of AI-VCR in IFN [7-9] (AI-Fuzzy VCN), we can make a realization of AI-VCR addition computing without overflowing as Fig.1. We definite some concepts from amongst the Initial Works: S_O=0 to be showed the SUM value S is not overflowing after addition computing, S_O=1 to be showed the SUM value S is overflowing after addition computing; SM is temporary variable for the SUM value. When S_O=0, the SUM figures (length of numbers) without changing; When S_O=1, the SUM figures (length of numbers) is in need of changing (increasing the length of number figures).

Sub-Prg1. When Valid number length of V_{Di} is longer than that length of V_D, namely, $V_D < V_{Di} \leq 2V_D$, this service program to be used in problem-solving of which expressing accumulate number A_i that as follows: $A_i = D_{ni} + D_{mi}$, $AM_i = A_i$, AM_i is temporary variable for the A_i value. All Figures of A_i would be saved at one-dimension variables of array in RAM(Read Access Memory) of computer.

Sub-Prg2. When SUM number S is overflowing, this service program to be used in problem-solving of which inlaying AI-VCR computation regularities of VCO, the typical FCN(Decimal numbers) would be changed into HCN(High Carrying Numbers), it is a kind of VCN which is input in RAM at one-dimension variables of array.

Sub-Prg3. When SUM number S is over, this service program to be used in problem-solving of which outputting the final SUM number S from the RAM, the HCN would be changed back into the typical FCN(Decimal numbers), it is very much in need of habituating to universal numbers for human beings.

In Fig.1, you are also to do many supplements else in the Initial Work, so that all of your sub-programs may be run correctly in later on, for example, public variables, VCR parameters, definite functions of AI, information of error prompts, information of inputting prompts, etc.

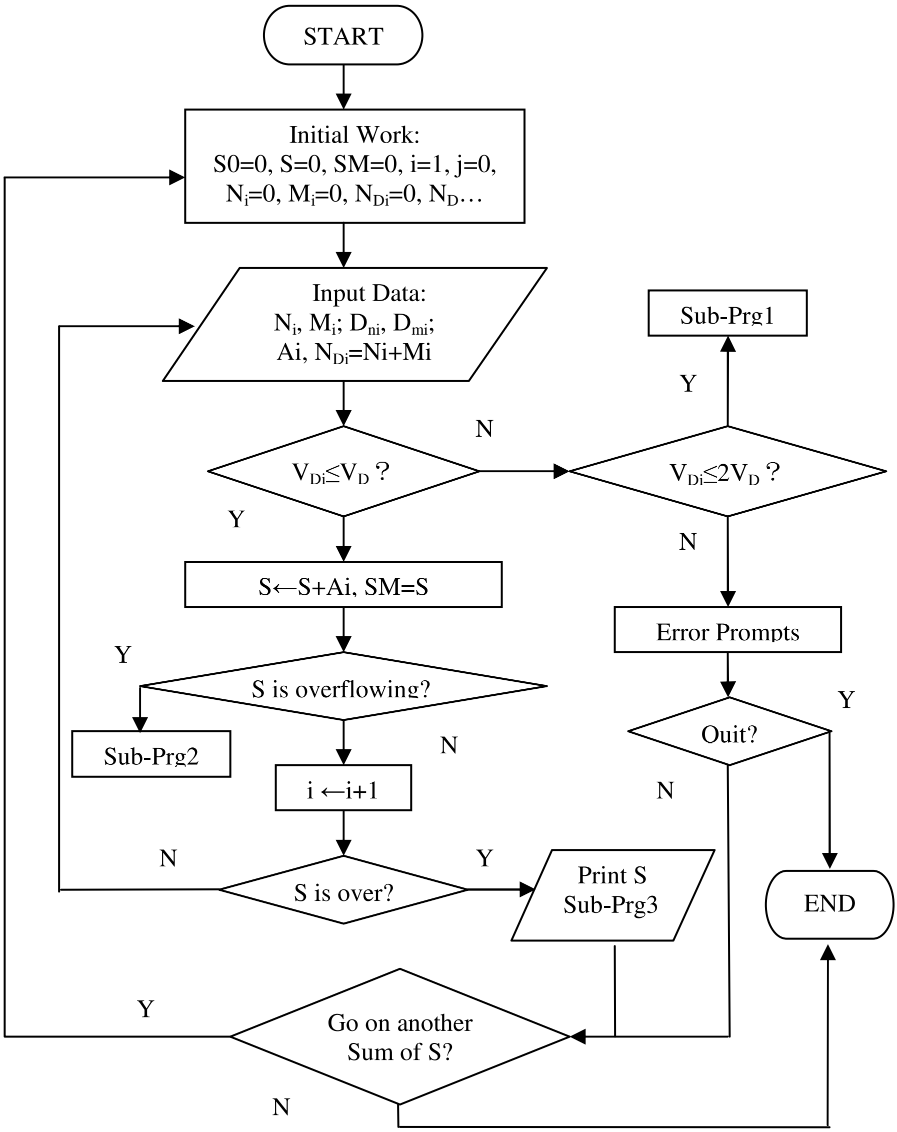


Fig. 1. Realization of AI-VCR addition computing without overflowing

5 Discussion of AI-VCR Addition Computing with VCO

In the above mentioned realization of addition computing, an addend and summand are the most typical FCN－Decimal numbers, owing to its universal research & application in the whole world for 15 centuries, it is in need of changing the HCN sum for addition computing back into the universal Decimal numbers; owing to the VCO of AI-VCR, the Number Value (NV) of an optional VCN wouldn't be overflowing from the sum data in RAM of computer, it looks like that, the flagon in hand of JIGONG monk in Chinese myth wouldn't be canned up with full of wine forever; owing to the Real VCN may be separately processed with both of Integer and Fraction, the Valid Figures to be used would be extended to $2V_D$, and the sum Valid Figures of CP would be limitless (even if it may surpass $2V_D$), accurate and without PM in computing of numbers, but the complexity of time for sum of PC would be added up in a great deal. For example, there are some VCN (Integer numbers or Real numbers) of 3-Figures such as $(516)_{FM2, FM1, FM0}$. Supposed that: V_D=8; FM_2, FM_1, FM_0 are to be some whole numbers (Integer numbers) which would be designed as you like, then you can compute lots of VCN accurately into decimal number [3] as follows:

$$I_{VCN} = (516)_{6,1234567891,9876543212}$$
$$= 5\times1234567891\times987654321+1\times9876543212+6\times9876543212^0$$
$$= 60,966,315,627,922,572,678D(DecimalNumber)$$
$$= 60966315627922572678(Total:20Figures = 2V_D + 4)\ ,$$

$$I_{VCN} = (516)_{8,123456,654321} = 5\times123456\times654321+1\times654321+6\times654321^0$$
$$= 403,899,921,207D(DecimalNumber)$$
$$= 403899921207(Total:12Figures = 2V_D - 4)\ ,$$

$$I_{VCN} = (516)_{6,10,9} = 5\times10\times9+1\times9+6\times9^0 \qquad (FM_2=6,\ FM_1=10,\ FM_0=9)$$
$$= 465D(DecimalNumber)$$
$$= 465(Total:3Figures = V_D - 5)\ ,$$

$$R_{VCN} = (51.6)_{6,10,8} = 5\times10+1\times10^0+6\times8^{-1} = 51.75D \text{ (4Figures=}V_D\text{-4)},$$

$$R_{VCN} = (51.6)_{10,2,12} = 5\times2+1\times2^0+6\times12^{-1} = 11.5D \text{ (3Figures=}V_D\text{-5)},$$

$$R_{VCN} = (51.6)_{10,2,9} = 5\times2+1\times2^0+6\times9^{-1} = 11.666\cdots D \text{ (Figures>2}V_D\text{)},$$

$$R_{VCN} = (5.16)_{10,2,12} = 5\times10^0+1\times2^{-1}+6\times12^{-1}\times2^{-1} = 5.75D \text{ (3Figures=}V_D\text{-5)},$$

$$R_{VCN} = (0.516)_{8,10,12} = 5\times8^{-1}+1\times10^{-1}\times8^{-1}+6\times12^{-1}\times10^{-1}\times8^{-1} = 0.64375D \text{(5Figures=}V_D\text{-3)}.$$

6 Conclusion

Because data word size (Bits of binary numbers) of an optional CPU of computer in the whole world is limited, it is in need of removing PM in all mathematics computation of numbers such as addition, subtraction, multiplication, division, etc. In this paper, there are 5 properties of realizing computation of AI-VCR addition as follows: (1) Sum of this addition isn't overflowing forever; (2) Sum valid figures of CP are limitless though valid figures of an addend & summand are limited; (3) The complexity of time for sum of PC is higher than that computation ones of universal FCN; (4) Sums of this addition are accurate and without PM, namely the errors of this addition are always equal to zero; (5) Valid Figures of addend and summand may be not over more than $2V_D$, but the Valid Figures of SUM after many times additions may be over more than $2V_D$.

The VCN for *n*-figures is a new and broad concept of numbers; and the FCN for *n*-figures is some of the special circumstances. Researches on the VCN for *n*-figures are to research into generality of the broad numbers. The application of VCN for *n*-figures is more widely and much more significance than that of FCN for *n*-figures.

The different additions of integer numbers (whole numbers), decimal/fraction numbers and real numbers (a number including integer numbers and decimal/fraction numbers) may be realized in different methods. For example, in integer numbers' addition, all figures in numbers of addend and summand must be eyes right (low figure to low figure); in decimal numbers' addition, all figures in numbers of addend and summand must be eyes left (high figure to high figure); in decimal/fraction numbers' addition, all figures in numbers of addend and summand must be eyes left (high figure to high figure); in real numbers' addition, all figures in numbers of addend and summand must be eyes middle (namely, integer figures to be eyes right, decimal or fraction figures to be eyes left).

In practice computing applications, additions given by this paper are only a few parts of making many important functions in the wide field of application. For example, the VCN of *n* figures still may be precisely used in counting Mental [1] Work, multi- expresses in a same value of AE function [10] , AI-searching [11] Technology, AI-express of diagram [12] in Computer Graphics, Numbers Compressive Techniques of AI, Multi-agent System (MAS) [13], Unified Theory of Information of Yixin Zhong [14], Synthesis Reasoning in Design [15], Extensive Researches of Numbers Theory [4,5,16] , Heapsort Algorithm [17], Multivariate Analysis [18], General Linear Model [19], Design and Analysis of Computer Algorithms [20], Networks Security of Information [21], etc.

For this reason, researches and applications on numbers wouldn't always be confined in a lane or in a narrow sphere of VCN such as binary numbers, the common decimal numbers, etc. The author's purpose of writing this paper is, to cast a brick to attract jade---offer a few common place remarks by way of introduction so that others may come up with valuable opinions, and to hope the more researcher will take much more interest in researches and applications on the VCN of *n*-figures later on, furthermore, the more achievements in their researches will come out.

Acknowledgements. The work was supported by the Natural Science Foundation of Fujian Province of China (A0640015) and the Universities Natural Science Fund for Department of Fujian Province of China (JA08246).

References

1. Wu, W.: Mechanization of mental work & modernization of science and technology in times of computer. The special invited paper in Progress of Artificial Intelligence in China. Press of Beijing University of Post & Telecom, Beijing (2003) (in Chinese)
2. Ma, Z.-m., Wang, Y.-s.: The master topics on numbers. Interviews Records of Zhiming Ma Academician of CAS (Chinese Academy of Science)/President of CMS (Chinese Mathematical Society)/Vice Chairman of IMU (International Mathematics Union), and Yusheng Wang Director of National Library of China, CCTV-10, Beijing, China, February 4 (2004)
3. Ye, Q.-s.: VCN & its role of engineering in human society. Engineering Sciences. Chinese Academy of Engineering, English Edition, Quarterly 6(1), 23–31 (2008)
4. Pan, C.-d., Pan, C.-b.: Basic numbers theory of analyzing. Science Press of China, Beijing (1997) (in Chinese)
5. Li, W.: Wang yuan topics on the guess of gold bach. Shandong Education Press of China, Jinan (1999) (in Chinese)
6. Ye, Q.-s.: A systematic information numbers technology with AI-property of VCR. Computer Systems & Applications (Chinese Academy of Sciences, Monthly) 18(12), 173–179 (2009) (in Chinese)
7. Ye, Q.-s.: Research & application of IFN in cypher science. Engineer-ing Sciences(Chinese Academy of Engineering, Monthly) 10(5), 51–58 (2008) (in Chinese)
8. Lu, Y.-p., Li, S.-z., Chen, S.-l., Zhou, C.-l., Guo, W.-z.: Particle swarm optimization based on adaptive diffusion and hybrid mutation. Journal of Software 18(11), 2740–2750 (2007) (in Chinese)
9. Zhang, T.-f., Ye, Q.-s.: A novel FPGA mathematics technology based AI-fuzzy VCN algorithm of automation. In: The Second International Workshop on Knowledge Discovery and Data mining (WKDD), vol. 1, pp. 323–327. IEEE CS, Moscow (2009)
10. Tu, X.-y. (ed.): Artificial emotion the special invited paper in Proceedings of CAAI-10, vol. 11, pp. 27–32. Beijing University of Post and Telecom Publishers, Beijing (2003 (in Chinese)
11. Shi, Z.-z. (ed.): Editor-in-Chief of Proc. of the third Pacific Rim International Conference (IFIP TC12/IEEE Computer Society/CCF/CAS) on Artificial Intelligence, PRICAI 1994. International Academic Publishers (IAP), Beijing (1994)
12. Qiusun, Y.: The numbers expression of 3D space visual solid reality pictures. Journal of Computer Aid Engineering (CAE) 6(4), 38–44 (1997) (in Chinese)
13. Cai, Z.-x.: Some advances in AI research, the special invited paper in Proceedings of CAAI-10, vol. 11, pp. 58–63. Beijing University of Post and Telecom Publishers, Beijing (2003) (in Chinese)

14. Zhong, Y.-x.: Intelligence Science: A Unified Theory of Information, Knowledge, Strategy and Its Execution, the special invited paper in Proceedings of CAAI-10, vol. 11, pp. 64–71. Beijing University of Post and Telecom. Publishers, Beijing (2003) (in Chinese)
15. Pan, Y.-h., Zhuang, Y.-t.: Towards AI of design & creation, the special invited paper in Proceedings of CAAI-10, vol. 11, pp. 16–27. Beijing University of Post and Telecom. Publishers, Beijing (2003) (in Chinese)
16. Cai, W., Yang, C.-y., Lin, W.-c.: Extension method of engineering. Science Press of China, Beijing (1997) (in Chinese)
17. Wang, X.-d., Wu, Y.-j.: An improved HEAPSORT algorithm with n log n - 0.788928n comparisons in the worst case. Journal of Computer Science and Technology 22(6), 898–903 (2007)
18. Wang, X.-d., Tian, J.: The estimation for interaction between the factors and the confidence intervals in multivariate analysis. Mathematics in Practice and Theory 39(10), 122–126 (2009)
19. Wang, X.-d., Tian, J.: The general linear model for measure of interaction between the exposure factors. Mathematics in Practice and Theory 40(7), 112–118 (2010)
20. Wang, X.-d.: The design and analysis of computer algorithms, 3rd edn. Publishing House of Electronics Industry of China, Beijing (2007)
21. B.-g. Lin, Networks security of information. Mechanics Industry Press of China, Beijing (2004) (in Chinese)

A New Face Recognition Algorithm Based on DCT and LBP

Pei-zhi Chen[1] and Shui-li Chen[2,⋆]

[1] School of Information Science and Technology, Xiamen University, Xiamen 361005, China
[2] School of Sciences, Jimei University, Xiamen 361021, China
e-mail: sgzx@jmu.edu.cn

Abstract. The paper proposes a new face recognition algorithm. Firstly, Discrete Cosine Transform (DCT) is conducted on an input face image. A few DCT coefficients on the left top corner are chosen as the global feature. At the same time, the image is divided into several parts. Local Binary Pattern (LBP)is conducted on each part and then LBP histogram sequences (Uniform LBP used) are accepted as the local feature. Secondly, the paper fusions the global and local feature using feature level fusion. Finally, Support Vector Machine (SVM) is adopted as the classifier and experiments are done on the ORL database. A result of 95.5% recognition rate and 5 ms time elapsed for each image is obtained, which shows the efficiency and practicability of the proposed algorithm and the correctness of feature level fusion. At last, the paper drops a conclusion that application of feature level fusion in face recognition will draw more and more attention in the future.

Keywords: Face recognition, Discrete Cosine Transform, Local Binary Pattern, Feature Level Fusion.

1 Introduction

Face recognition is a typical research problem on model recognition, image analysis and understanding. It has high research value. At the same time, face recognition has high application value in many fields such as public security, banking and so on. Therefore, face recognition is a focus in both academia and industry.

DCT is an important data compression technique. It has been widely used in JPEG and MPEG standard. ZIAD M. HAFED et al. first introduced DCT into the face recognition field [1]. DCT has special high efficiency and data independence. LBP is a texture representing approach. It has showed high robustness and efficiency in the face recognition field [2].

⋆ Corresponding author.

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 811–818.
springerlink.com

However, performance of single DCT or LBP is not that perfect. Kittler et al. developed a common theoretical framework for combining classifiers and four fusion rules [3]. Features in different models were abstracted and classified independently. Several recognition matrixes were obtained. Final decision according to fusion rules was made. In the face recognition experiment, three different features of front face images, profile face images and voice were combined. The result successfully verified the theoretical framework and the nice classifying performance of the sum rule in the paper. Afterwards, decision level fusion attracted more and more researchers' attention. It has been widely used in face recognition fusion approaches [4, 5, 6], and a nice recognition result has been achieved.

But decision level fusion is very complex due to its muti-classifiers, which limit its application in practical use. As a starting point, our paper intends to investigate feature level fusion which only needs one classifier. Because several different features of input data were collected, the classifier has more sufficient matching scores or decision results [7]. As a result, we believe that feature level fusion will perform better than decision level fusion.

The paper combines DCT and LBP feature using feature level fusion. Experiment results show that our algorithm can achieve high recognition rate and fulfill real-time application requirement.

SVM using RBF kernel function is accepted as the classifier in the paper.

2 Face Representation

The overall framework of the proposed representation algorithm based on feature level fusion of LBP and DCT is illustrated in Fig.1. In our algorithm,

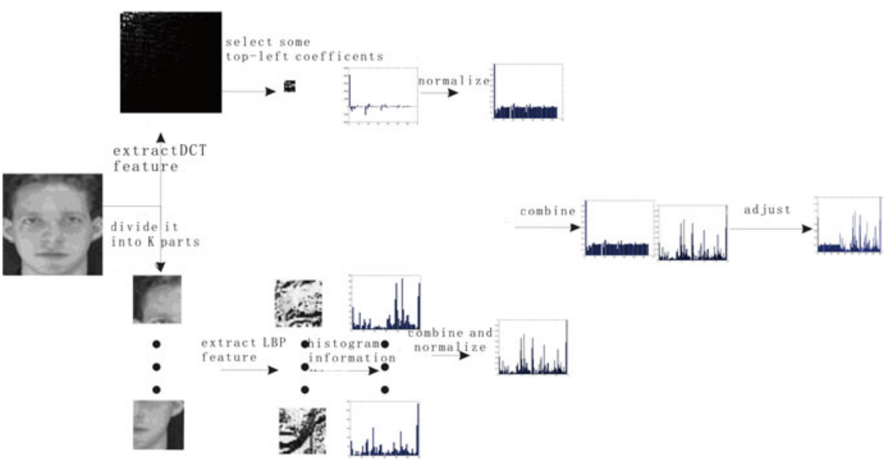


Fig. 1. Steps of the proposed algorithm in the paper

the feature of an input face image can be obtained by the following procedures:(1)DCT is conducted on an input face image and a few DCT coefficients on the top left corner are selected as the global feature;(2)At the same time, the image is divided into k non-overlapping regions and then transformed to get LBP histogram sequences as the local feature.(3)Normalize the global and local feature and adjust them according to the strategy proposed in the paper.(4)Combine the adjusted features as the final features of the image.The procedures can be instructed in Fig.1.

2.1 DCT

DCT is an orthogonal transformation proposed by N.Ahmed et al. in 1974. It is always considered to be the most optimal transformation method on voice and image signal processing. In the image processing field, to conduct two dimensional DCT is to separate high frequency and low frequency information.

Given an input image A sized M*N, its two dimensional DCT result B is obtained by the following equation:

$$B_{pq} = \alpha_p \alpha_q \sum_{m=1}^{M} \sum_{n=1}^{N} cos\frac{\pi(2m+1)p}{2M} cos\frac{\pi(2n+1)q}{2N} \quad 0 \le p \le M-1 \qquad 0 \le q \le N-1 \tag{1}$$

$$\alpha_p = \begin{cases} \sqrt{\frac{1}{M}} & p = 0 \\ \sqrt{\frac{2}{M}} & 1 \le p \le M-1 \end{cases} \tag{2}$$

$$\alpha_q = \begin{cases} \sqrt{\frac{1}{N}} & q = 0 \\ \sqrt{\frac{2}{N}} & 1 \le q \le N-1 \end{cases} \tag{3}$$

Obviously, the size of the result coefficient matrix B is M*N. Therefore, pure DCT transform can not play the role of data reduction.

As is depicted in Fig.2, large DCT coefficients concentrate on the top left corner. These coefficients show the low frequency information of the face. As

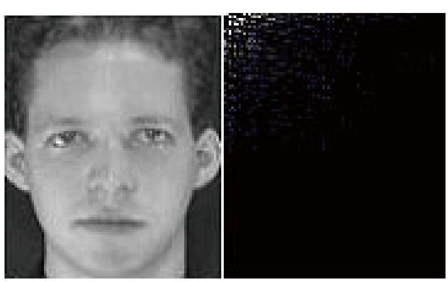


Fig. 2. DCT results

a result, these top left coefficients can be accepted as the representation of the face. According to results in the paper [1], choice of 8*8 sized region on the left-lop shows the best recognition effect. So the same region size 8*8 is adopted in our paper as the global feature of the face.

2.2 LBP

LBP was introduced by Ojala in 1996. It is a powerful method of texture description based on statistical analysis and shows its practical use in texture description. LBP was applied in the face recognition field in 2004 and showed nice discriminability and fast recognition speed [2]. The operator labels the pixels of an image by thresholding the 3*3 neighborhood of each pixel with the center pixel and considering the result as a binary number. The LBP result can be calculated as follows:

$$LBP_{P,R} = \textstyle\sum_{p=0}^{P-1} s(g_p - g_c) * 2^p \tag{4}$$

Where g_c is the center pixel value and g_p is one of the neighborhoods around the center with the radius R, P is the whole neighborhood number.

$$s(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{5}$$

For example,$LBP_{8,1}$represented that radius R is 1 and there are P(P=8)neighborhoods all together, as is depicted in Fig.3:

Obviously, for an input image sized M*N, the LBP image's size is (M-1)*(N-1).

The original LBP with P neighborhoods has 2^p different binary pattern. For example, given P=8, LBP operator gets 256 different binary patterns. Therefore, Ojala made a definition of Uniform LBP noted $LBP^u_{P,R}$, for those Patterns which contain at most two bitwise transitions from 0 to 1 or vice versa when the binary strings are considered circular. For example, 00000000, 11110000 and 01110000 are uniform patterns. For those patterns that are not uniform patterns, we fall them under one special pattern, for example 11110010 can be used to represent them.Therefore, when P=8, $LBP^u_{P,R}$ has

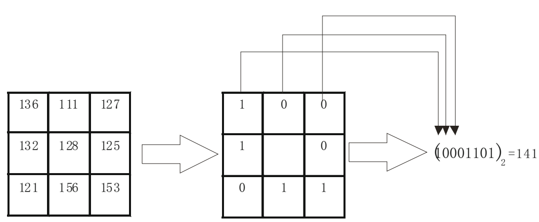


Fig. 3. LBP

all together 58+1 different patterns. In our paper, $LBP^{u}_{P,R}$is adopted for the image texture and them histogram sequence information of the texture image is regarded as the feature.

The histogram sequence of the LBP texture image f(x,y)can be obtained as follows:

$$h_i = \sum I[f(x,y) = i] \qquad \textbf{i is one of the uniform patterns} \tag{6}$$

$$I(A) = \begin{cases} 1 & \textbf{A is true} \\ 0 & \textbf{A is false} \end{cases} \tag{7}$$

2.3 Feature Level Fusion

A lot of physiology and psychology research shows that global feature and local feature play different roles on face representation and recognition. The global feature reflects the face's whole property while the local feature reflects the detail change on the face.

Given a face imageA_{m*n}whose DCT coefficient is $D_1 D_2 \cdots D_{m1}$ and LBP histogram sequence $L_1^1 L_2^1 \cdots L_{59}^1 L_1^2 L_2^2 \cdots L_{59}^2 \cdots L_1^k L_2^k \cdots L_{59}^k$,the procedure of feature level fusion of DCT coefficient is showed as follow:

(1) Data normalization

DCT coefficient and LBP histogram sequence are separately normalized to [0,1]; The normalized DCT coefficients are noted

$$D_1^n D_2^n \cdots D_{m1}^n$$

While the normalized LBP histogram sequences are noted

$$L_1^{1,n} L_2^{1,n} \cdots L_{59}^{1,n} L_1^{2,n} L_2^{2,n} \cdots L_{59}^{2,n} \cdots L_1^{k,n} L_2^{k,n} \cdots L_{59}^{k,n}$$

(2) Compute adjust weight

Given the mean value of the DCT coefficient and LBP histogram sequence of the training set respectively $mean_{dct}$ and $mean_{lbp}$, we can obtain the LBP weight is

$$\frac{mean_{dct}}{mean_{lbp}}$$

(3) Change the LBP feature

Change the LBP histogram sequence of the training set and the test set as follows:

$$L_i^{j,c} = \frac{mean_{dct}}{mean_{lbp}} * L_i^{j,n}$$

Where i=1, 2, $\cdots$, 59   j=1, 2,$\cdots$,k

(4) Combine feature

Combine the treated DCT coefficient and LBP histogram sequence and finally get the feature of A_{m*n}as follow:

$$D_1^n D_2^n \cdots D_{m1}^n L_1^{1,c} L_2^{1,c} \cdots L_{59}^{1,c} L_1^{2,c} L_2^{2,c} \cdots L_{59}^{2,c} \cdots L_1^{k,c} L_2^{k,c} \cdots L_{59}^{k,c}$$

3 Experiment

In this section, we will demonstrate the robustness of the combination of DCT and LBP in feature level fusion .We choose the AT&T face database[8] (ORL database) in our experiments. The subset of ORL database includes 400 images of 40 individuals (each individual has 10 images), for some subjects, the images were taken at different times, varying the lighting, facial expressions (open / closed eyes, smiling / not smiling) and facial details (glasses / no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position (with tolerance for some side movement).The size of each image is 92*112 pixels.In our paper the size is changed to 64*64 pixels.Ten images of one person is showed in Fig.4.

Fig. 4. Ten face images in ORL database

We perform our algorithm on a MS Windows XP PC with dual Intel Xeon 2.8-GHz CPUs, 4.0 Gbytes of RAM, using Matlab Programming.

In this paper, all the experiments are done using LIBSVM [9] developed by Lin Zhiren . 8 * 8 region of the DCT coefficients of the left top corner is chosen, and the image is divided into 4 sub-blocks to abstract LBP features. For each individual, the former five images are chosen for training , while the rear 5 images for testing. In the LIBSVM, c=2.50, g=0.16.

The recognition result is showed in Tab.1.

The results show that, our algorithm performed better in recognition rate then DCT or LBP separately due to the combination of global feature and local feature in our algorithm.

The training and testing time of our algorithm is showed in Tab.2

Table 1. Comparation of recognition results

Method	Recognition rate (%)
DCT+SVM	84
LBP+SVM	90.5
DCT+LBP+SVM	95.5

Table 2. Time used in our proposed algorithm

Items	Time (s)
training	1.121361
testing	1.106737

Testing time for each face image is $1.106737s/200 \approx 5ms$ As can be seen from our experiments, each face image costs 5ms and the recognition rate is 95.5% as long as we perform training in advance. The results meet the practical application request. Using support vector machine as the classifier, which can be implemented by hardware method, makes our algorithm more suitable for practical application in engineering.

4 Conclusions

Experiment results show that the feature fusion level of DCT and LBP algorithm are robust to change in expression,pose,illumination. Therefore, we can drop an conclusion that our algorithm is suitable for practical application request.

At the same time, the paper shows the effectiveness of feature level fusion. That is to say, the proposed algorithm has certain extendness. It is believed that feature level fusion will be researchable and draw more and more attention in the face recognition field.

Acknowledgements. The project is supported by the National Natural Science Foundation of Fujian Province, PRC (2008J0024); The key project of education department of Fujian province, PRC(JK2009017); The key project of institutes serving for the economic zone on the western coast of the Taiwan strait (2009(8)); The National Natural Science Foundation of Xiamen City, PRC (3502Z2093018).

References

1. Hafed, Z.M., Levine, M.D.: Face Recognition Using the Discrete Cosine Transform. International Journal of Computer Vision 43, 167–188 (2004)
2. Ahonen, T., Hadid, A., Pietikainen, M.: Face recognition with local binary patterns. In: Pajdla, T., Matas, J(G.) (eds.) ECCV 2004. LNCS, vol. 3021, pp. 469–481. Springer, Heidelberg (2004)
3. Kittler, J., Hatef, M., Duin, R.P.W., et al.: On Combining Classifiers. IEEE Transactions on Pattern Analysis and Machine Intelligence 20, 226–239 (1998)
4. Wenchao, Shiguang, Z., Gao, S.W., et al.: Information Fusion in Face Identification. In: International Conference on Pattern Recognition, vol. 3, pp. 950–953 (2004)

5. Lu, X., Wangy, Y., Jain, A.K.: Combining classifiers for face recognition. In: International Conference on Multimedia and Expo., vol. 3, pp. 13–16 (2003)
6. Liu, Z., Liu, C.: Fusion of the complementary Discrete Cosine Features in the YIQ color space for face recognition. Computer Vision and Image Understanding 111, 249–262 (2008)
7. Aroussi, E., Amine, M., Ghouzali, A., et al.: Combining DCT and LBP Feature Sets For Effcient Face Recognition. In: International Conference on Information and Communication Technologies: From Theory to Applications, pp. 1–6 (2008)
8. ORL face database (1994), http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html
9. LIBSVM (2000), http://www.csie.ntu.edu.tw/~cjlin/libsvm/

An Building Edge Detection Method Using Fuzzy SVM

Yun-dong Wu[1,2], Wen An[1,3], Qiang Zhang[3], and Shui-li Chen[1]

[1] Research Centre of Image Information Engineering ,Jimei Univ., 185 Yinjiang Rd., Xiamen, Fujian, China
[2] School of Sciences, Jimei Univ., 185 Yinjiang Rd., Xiamen, Fujian, China
[3] Institute of Surveying and Mapping, Information Engineering Univ., 66 Longhai Rd., Zheng zhou, Henan, China
e-mail: wuyundong@yahoo.com.cn

Abstract. According to the fuzzy edge information, this paper presents a FSVM(Fuzzy Support Vector Machine) edge detection algorithm has its own advantages. First, it calculates each point's 3 × 3 window of the four directions of the difference as gradient and it regards gradient value and four differences and gradient angle as edge point feature. It utilize gauss RBF kernel as kernel function. Next, it constructs fuzzy edge point's membership function. Finally, it uses FVM to do edge detection. Experiments show that our FSVM(Fuzzy Support Vector Machine) edge detection algorithm is feasible.

Keywords: FSVM, membership function, image processing, edge detection.

1 Introduction

Edge detection is the most fundamental problem in image processing. Over the years, many methods have been proposed for detecting edges. Some of the earlier methods, such as Sobel, Prewitt and Robert and Canny [1]. Those methods base on edge point's gray gradient [2], set threshold to find edge point algorithm.Canny uses dual-threshold to detect edge. This method obtains favorable edge results.

In recent years, there is another novel edge detection method, which considers edge detection as two classification problem that is edge point and non-edge point. In paper [3], Wei G et al introduces SVM to detect edge. Support vector machines (SVM) introduced by Vapnik are based onstatistical learning theory. An SVM first maps the edge point input feature into a high-dimensional feature space and finds a separating hyperplane that maximizes the margin between two classes in this space .SVM is a powerful tools for solving two classification problems. But edge point and non-edge point are not exactly assigned to one of those classes. Some date points are more important. Some date points such as noise are less meaningful .SVM can't consider this situation. In paper [4], Hong et al apply a fuzzy membership to each input point of SVM and reformulate SVM into fuzzy SVM (FSVM) so that different input points can make different contribution to the learning of decision surface. This method can reduce the effect of noises in data

B.-Y. Cao et al. (Eds.): Quantitative Logic and Soft Computing 2010, AISC 82, pp. 819–826.
springerlink.com

points. In this paper, we retain the traditional edge detection's gradient information and use SVM powerful two classification ability, introduce the fuzzy membership functions, handle the transition between edge points and non-edge points problem. We introduce Fuzzy SVM to detect edge points. We will not introduce Fuzzy logic in detail; it can be seen in [5]. Finally, we test our method and compare to Sobel and SVM edge detection.

2 Construct Edge Point Feature

From eqs.(1) to eqs.(6) we give the formula of each pixel's 3×3 neighborhood as well as the difference of 0^o , 90^o, 45^o, 135^o four directions, gradient values and direction.

0^o direction difference:

$$p_1[i,j] = I[i+1,j] - I[i-1,j] \ , \tag{1}$$

90^o direction difference:

$$p_2[i,j] = I[i,j+1] - I[i,j-1] \ , \tag{2}$$

45^o direction difference:

$$p_3[i,j] = I[i-1,j+1] - I[i+1,j-1] \ , \tag{3}$$

135^o direction difference:

$$p_4[i,j] = I[i+1,j+1] - I[i-1,j-1] \ . \tag{4}$$

Gradient value and direction can be calculated as:

$$m[i,j] = \sqrt{p_1[i,j]^2 + p_2[i,j]^2 + p_3[i,j]^2 + p_4[i,j]^2} \ , \tag{5}$$

$$\theta(i,j) = \arctan\left(\frac{p_1[i,j]}{p_2[i,j]}\right) \ . \tag{6}$$

Normalized gradient value:

$$m[i,j]' = \frac{m[i,j] - M_{\min}}{M_{\max} - M_{\min}} \ . \tag{7}$$

According to traditional edge detections' result, we have good reason believe that edge points of information is focus on gray and gradient of each pixel's 3×3 neighborhood, and building edge points' gray and gradient have strong directions. So we construct each training sample point feature as six-dimension

$$x_i = (p_1, p_2, p_3, p_4, m', \theta) \ .$$

3 SVM

SVM (support vector machine) is a supervised learning technique from the field of machine learning applicable to both classification and regression which is based on the principle of Structural Risk Minimization [6] . The algorithm can be summarized as mapping the data into a very high dimensional feature space using an appropriate kernel function, constructing an optimal separating hyper-plane, and then trying to separate the mapped vectors from the origin with maximum margin. The optimal separating hyper-plane is each class edge. We will not introduce it in detail; it can be seen in [6]. Its idea is as follows. $S = \{x_i, y_i\}$ Each training point $x_i \in R^m$ has a label $y_i = +1$ or $y_i = -1$. In this paper, where $x_i = (p_1, p_2, p_3, p_4, m', \theta)^T$ and edge point $y_i = +1$, non-edge point $y_i = -1$. Next, We wish to find the hyper-plane
$(w \cdot x) + b = 0$ .Edge training point hyper-plane:

$$w^T x_i + b \geq 1, (i = 1, \cdots, n_1) \quad (8)$$

Non-edge point hyper-plane:

$$w^T x_i + b \leq -1, (i = n_1, \cdots, n) \quad (9)$$

The optimal hyper-plane can be obtained by solving the following convex quadratic optimization problem

$$\begin{cases} \min \quad \phi(w) = \frac{1}{2} w^T w \\ s.t \quad y_i \left[w^T x_i + b \right] - 1 \geq 0 \end{cases} \quad (10)$$

We construct the Lagragian and transform into its dual problem.

$$L(w,b,a) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{n} a_i \left\{ y_i[(w^T x_i) + b] - 1 \right\}, \quad (11)$$

$$\begin{cases} \max \phi(a) = \sum_{i=1}^{n} a_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j y_i y_j x_i^T x_j \\ s.t \quad \sum_{i=1}^{n} a_i \, y_i = 0; 0 \leq a_i \leq C \end{cases}, \quad (12)$$

where a_i is lagragian parameter, can be calculate by Lagragian.

At last the optimal classification function can be described as

$$f(x) = \operatorname{sgn}((w_0 \cdot x) + b_0), \quad (13)$$

where w_0 , b_0 :

$$w_0 = \sum_i^n a_i y_i x_i, \tag{14}$$

$$b_0 = -\frac{1}{2} w_0 \left(x_i \cdot x \right). \tag{15}$$

In this paper, each training point feature is six-dimension. Our separable problem is nonlinear. So we need kernel function to solve this problem. In this paper, our kernel function is Gaussian kernel function. The optimal classification function can be described as

$$f(x) = \mathrm{sgn}(\sum_{j=1}^{l} a_j^* y_j K(x_i, x) + b_0), \tag{16}$$

where Gaussian kernel function:

$$K\left(x, x'\right) = \exp\{-\frac{\|x - x'\|}{\sigma^2}\}. \tag{17}$$

4 Fuzzy SVM

SVM is a commonly used method in statistical learning theory. SVM can only solve one or the other classification. In fact some training points are not so absolute, which certain some degrees of ambiguity, some data points are more important, and some data points are less important, so the machine should take those situation into account. SVM lack this kind of ability, so they reformulate SVM into Fuzzy SVM [4,5,7] (FSVM) .Its idea is as follows. Given a set S_f of labeled training points with fuzzy memberships $S_f = \{x_i, y_i, s_i\}$ .Each training point $x_i \in R^m$ has a label $y_i = +1$ or $y_i = -1$, where $x_i = \left(p_1, p_2, p_3, p_4, m', \theta\right)^T$ and edge point $y_i = +1$, non-edge point $y_i = -1$, and a fuzzy memberships $0 \le s_i \le 1$.

According to the characteristics of normalized edge gradient value, in this paper the proposed S fuzzy membership can be described as follows:

$$S\left(m', a, b, c\right) = \begin{cases} 0, & m' < a \\ \dfrac{(m'-a)^2}{(b-a)(c-a)}, & a \le m' < b \\ 1 - \dfrac{(m'-c)^2}{(c-b)(c-a)}, & b \le m' < c \\ 1, & c \le m' \end{cases}, \tag{18}$$

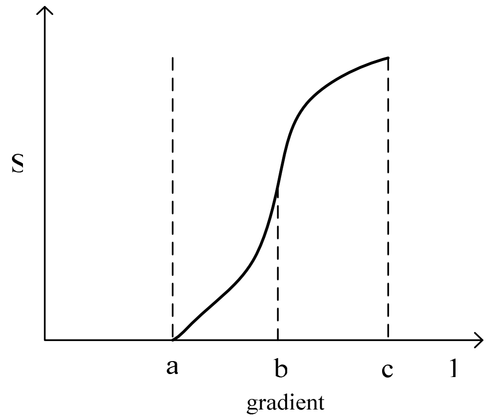


Fig. 1. S fuzzy membership

where a、b、c determine the shape of S fuzzy membership function. Fuzzy value b is usually the middle of [a, c] .Figure.1 shows S membership function curve image.

S fuzzy membership function can fuzzy the process which from edge to non-edge. So there are two fuzzy states that the possibility of non-edge and the possibility of edge.

The optimal hyper-plane problem is then regarded as solution to

$$\begin{cases} \min \quad \phi(\omega,\xi,s)=\frac{1}{2}\omega^T\omega+C\sum_{i=1}^{n}s_i^m\xi_i \\ s.t \quad y_i\left[\omega\times\phi(x_i)+b\right]\geq 1-\xi_i \end{cases}, \tag{19}$$

where C＞0 is a penalty factor for errors. Tuning this parameter can make balance between margin maximization and classification violation. The above problem is an optimization problem .In order to solve this problem, we construct the Lagragian and transform into its dual problem.

$$\begin{cases} \max\phi(a)=\sum_{i=1}^{n}a_i-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}a_ia_jy_iy_jx_i^Tx_j \\ s.t \quad \sum_{i=1}^{n}a_i\,y_i=0;0\leq a_i\leq s_iC \end{cases}. \tag{20}$$

According to the theory of FSVM, the nonlinear separable problem can be solved by using the kernel function.

$$\begin{cases} \max\phi(a)=\sum_{i=1}^{n}a_i-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}a_ia_jy_iy_jK(x_i,x) \\ s.t \quad \sum_{i=1}^{n}a_i\,y_i=0;0\leq a_i\leq s_iC \end{cases}. \tag{21}$$

At last the optimal classification function can be described as

$$f(x) = \mathrm{sgn}(\sum_{j=1}^{l} a_j^* y_j K(x_i, x) + b_0) \; . \tag{22}$$

5 Experiments and Results

5.1 Image Edge Detection Steps Base on FSVM

Our method is composed of 4 steps, and is described as follows and shown in Figure.2:

1) Using improved Canny[8] detect edge. We take edge points as edge training points (x_i, y_i), and manually select the non-edge training points, calculate each training point's feature
2) According to the method which is described in formula (21), finding out the membership S_i for each training point.
3) Constructing the FSVM by introducing memberships (which is obtained form step 2) to the SVM.
4) Using the FSVM to detect the gray image edge.

In this paper, comparing to FSVM edge detection, SVM edge detection don't need fuzzy memberships, other steps are the same, such as edge point feature and training sample and so on.

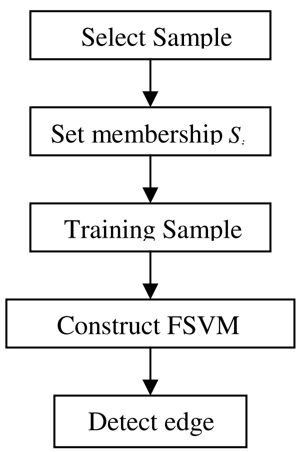


Fig. 2. FSVM edge detect steps

5.2 Simulation and Results

Simulation environment is Windows XP and MATLAB R2009a software, hardware environment is T660, 2G of main memory. The proposed algorithm is applied on test picture which is 500×319 pixels acquired with UAV platform in Guanqiao Town.

In this paper, we find the good training effect when (C, σ, a, c) and (C, σ) is equal to (4500, 0.2, 0.2, 0.8), (4500, 0.2), where σ is Gaussian K ' parameter, C is penalty factor for errors. The results are compared with the other edge detection method such as Sobel, SVM. The experiment results are shown in Figure.3, Parameters and Run-time are shown in Table 1.

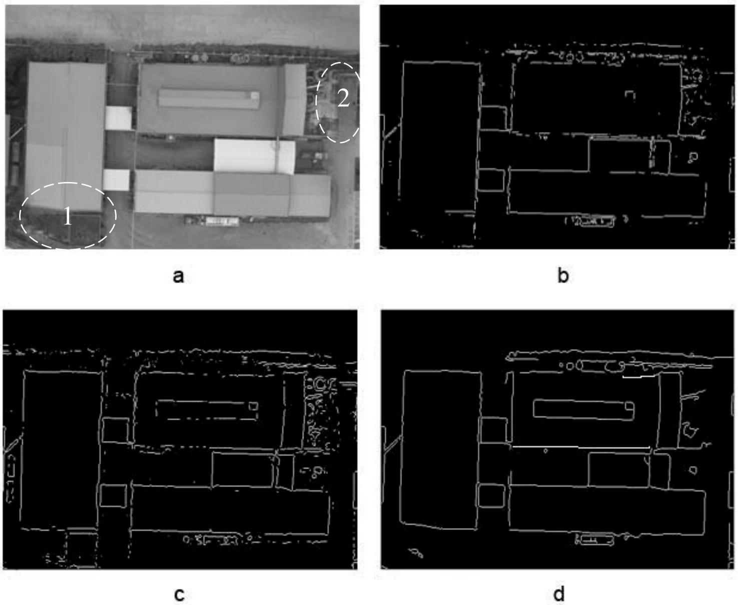


Fig. 3. a)Main image, b) Sobel results, c)SVM result, d)FSVM results

Table 1. Run-times and parameter

	parameter	Run-times
Sobel	Thre=0.05	Time=3.56s
SVM	σ=0.2 C=2500	Time=5.60s
FSVM	a=0.2 c=0.8 σ=0.2 C=2500	Time=7.30s

Comparing region 1 of Figure 3's edge detect result , Sobel edge detect result (b) lost a lot of edge, SVM (c) is better than Sobel, FSVM edge detect result (d) is clear and complete, so FSVM can better maintain the integrity of edge features. From figure.3, edge detect results in region 2 show that FSVM edge detector can better avoid noise where building outside.

6 Conclusion

Throughout this paper, we introduce FSVM to detect building edge. The FSVM edge detector includes appropriate defined fuzzy membership function and decided about pixel classification as edge or non-edge. Experimental results shown this method extract more integrity of edge and avoid more noise than Sobel. And it is shown that introduce fuzzy memberships to improve SVM is useful. We will do more research in the relationship between Kernel function and building edge detection in future.

Acknowledgements. The work was supported by The key project of institutes serving for the economic zone on the western coast of the Taiwan strait (2009(8)); The key project of education department of Fujian province, PRC (JK2009017); The National Natural Science Foundation of Xiamen City, PRC (3502Z2093018).

References

1. Canny, J.: A Computational Approach to Edge Detection. IEEE Transactions on Pattern Analysis And Machine Intelligence PAMI 8(6), 679–699 (1986)
2. Lily, R.L., Carl, G.L.: Competitive fuzzy edge detection. Applied Soft Computing (3), 123–137 (2003)
3. Wei, G., Zhenhong, J.: Improved Edge Detection Based on LS-SVM: Fourth International Conference on Natural Computation, pp. 86–90 (2008)
4. Hong, X.: Feature Selection based on Fuzzy SVM. Fifth International Conference on Fuzzy Systems and Knowledge Discovery, 586–589 (2008)
5. Chen, S.-l., et al.: Fuzzy Sets Logic and its Applications. Science Press (2005)
6. Vapnik, V.: The Nature of Statistical Learning Theory. Springer, Berlin (1995)
7. Inoue, T., Abe, S.: Fuzzy Support Vector Machines for Pattern Classification: Neural Networks International Joint Conference, pp. 1449–1454 (2001)
8. Wang, Z., et al.: An Adaptive Edge-detection Method Based on Canny Algorithm. Journal of Image and Graphics 9(8), 958–962 (2004)

Author Index