ANALYSIS, DESIGN AND EVALUATION OF MAN-MACHINE SYSTEMS 1995

A Postprint volume from the Sixth IFAC/IFIP/IFORS/IEA Symposium, Cambridge, Massachusetts, USA, 27 - 29 June 1995

Edited by

T.B. SHERIDAN
Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

(In two volumes)

Published for the

INTERNATIONAL FEDERATION OF AUTOMATIC CONTROL

by

PERGAMON
An Imprint of Elsevier Science

UK	Elsevier Science Ltd, The Boulevard, Langford Lane, Kidlington, Oxford, OX5 1GB, UK
USA	Elsevier Science Inc., 660 White Plains Road, Tarrytown, New York 10591-5153, USA
JAPAN	Elsevier Science Japan, Tsunashima Building Annex, 3-20-12 Yushima, Bunkyo-ku, Tokyo 113, Japan

First edition 1995

Library of Congress Cataloging in Publication Data

A catalogue record for this book is available from the Library of Congress

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN 0-08-042370 1

This volume was reproduced by means of the photo-offset process using the manuscripts supplied by the authors of the different papers. The manuscripts have been typed using different typewriters and typefaces. The lay-out, figures and tables of some papers did not agree completely with the standard requirements: consequently the reproduction does not display complete uniformity. To ensure rapid publication this discrepancy could not be changed: nor could the English be checked completely. Therefore, the readers are asked to excuse any deficiencies of this publication which may be due to the above mentioned reasons.

The Editor

Printed in Great Britain

6th IFAC/IFIP/IFORS/IEA

SYMPOSIUM ON ANALYSIS, DESIGN, AND EVALUATION OF MAN-MACHINE SYSTEMS

Hosted by

Department of Mechanical Engineering and
Department of Aeronautics and Astronautics
Massachusetts Institute of Technology
Cambridge, MA USA
June 27-29, 1995

Sponsored by:
International Federation of Automatic Control (IFAC)
Technical Committee on Systems Engineering
Technical Committee on Social Effects of Automation
American Automatic Control Council

Co-Sponsored by:
International Federation for Information Processing (IFIP)
International Federation of Operational Research Societies (IFORS)
International Ergonomics Association (IEA)

International Program Committee (IPC):
Chairman: T.B. Sheridan (USA)

J.L. Alty (GB)
A.B. Aune (N)
S. Baron (USA)
P. Brodner (D)
P.C. Cacciabue (I)
W.J. Edwards (Aus)
J.L. Encarnacao (D)
R. Genser (A)
G. Giuda (I)
B.S. Hu (PRC)
R. Jackson (USA)
G. Johannsen (D)
K. Kawai (J)
V. deKeyser (B)
A.H. Levis (USA)
M. Lind (DK)
B. Liu (PRC)
N. Malvache (F)
L. Martensson (S)
P. Milgram (CAN)
C. Mitchell (USA)
N. Moray (USA)
L. Motus (ES)
L. Nemes (AUS)
W.K. Oxenaar (NL)
J. Ranta (SF)
J.E. Rijnsdorp (NL)
W.B. Rouse (USA)
A.P. Sage (USA)
G. Salvendy (USA)
H.G. Stassen (N)
S.Q. Su (PRC)
B.G. Tamm (ES)
H. Tamura (J)
T. Vamos (H)

National Organising Committee (NOC)
Chairman: R. John Hansman, MIT

History of the IFAC Symposium on Man-Machine Systems

This is the sixth in a series of (usually) triennial symposia on man-machine systems under IFAC sponsorship. It is the first of the series to be held in the United States. The sequence has been:

1982 Baden-Baden, Federal Republic of Germany
1985 Varese, Italy
1988 Oulu, Finland
1989 Xi'an, China
1992 The Hague, The Netherlands
1995 Cambridge, Massachusetts, USA

A NEW HUMAN-COMPUTER-INTERFACE FOR HIGH-SPEED-MAGLEV TRAIN TRAFFIC SUPERVISION

Jens-Olaf Müller, Eckehard Schnieder

Institut für Regelungs- und Automatisierungstechnik, Langer Kamp 8, D-38106 Braunschweig,
Tel.: +49 531 391 3317 / FAX: +49 531 391 5197, e-mail: jom@ifra.ing.tu-bs.de

Abstract: In modern high speed train traffic high capacity and acceptance of the transport system becomes more and more important. Human Computer Interfaces nowdays used in train traffic control are developed historically and do not fit any longer as the real traffic situation is difficult to perceive by the dispatcher. In this article, weaknesses of existing interfaces are shown. A completely new designed interface is proposed consisting of completely new designed diagrams helping the dispatcher to optimally perceive the traffic situation. It is shown that interface design can no longer be limited to a simple screen design, but the whole working area has to be considered. For process control applications new diagram design ideas considering human mental work-load and human mental capacity should be invented.

Keywords: human-machine interface, process control, train control, traffic control, visual pattern recognition

1. INTRODUCTION

In our society transportation has become an important factor of increased life quality on the one hand and an important economic factor on the other. Time is money. Human beings want to travel longer and longer distances in less and less time. In european countries like Germany, a high standard of life means high mobility. As car traffic density increases this transport medium begins to collapse. Unfortunately, space to broad highways do not exist. So alternatives have to be found. They are seen in the german high speed train ICE and the maglev train TRANSRAPID (Schnieder, 1983). To make these transport systems acceptable by the passengers and profitable for the owner, a high traffic density (=high capacity), high speed, high transport comfort and no delays must be assured. High speed trains are controled by a dispatcher sitting in a control tower. He needs to have an overview of the whole traffic situation. In case of occuring future conflicts on the railway net, he has to find alternatives just in time. A maglev train with a velocity of about 500 km/h passes 30 km in about 4 minutes. So the decision interval for the dispatcher is very short.

HCIs nowadays used in train traffic control have developped historically. For train velocities and densities existing today and in future they do no longer a good job. The following ideas introduced to the construction of those interfaces will lead up to a new era of train traffic control.

2. EXISTING INTERFACES

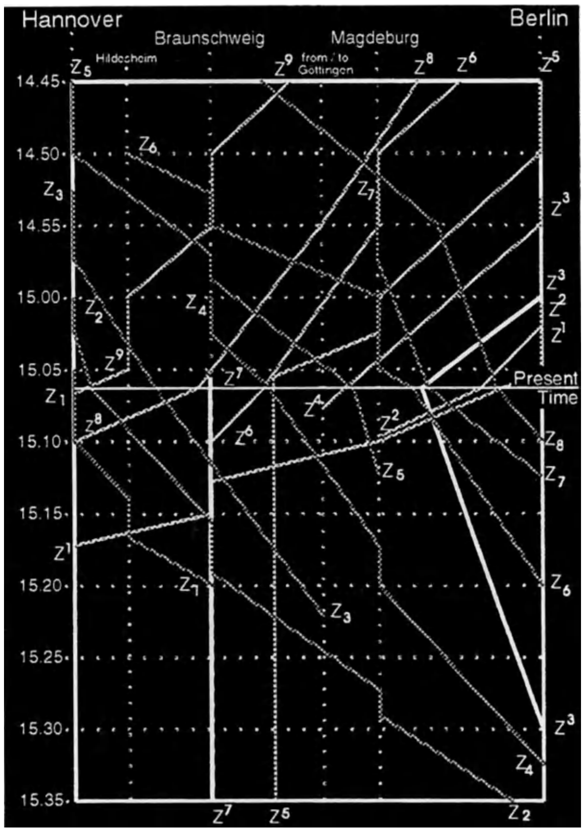


Fig. 1: Time-travel diagram

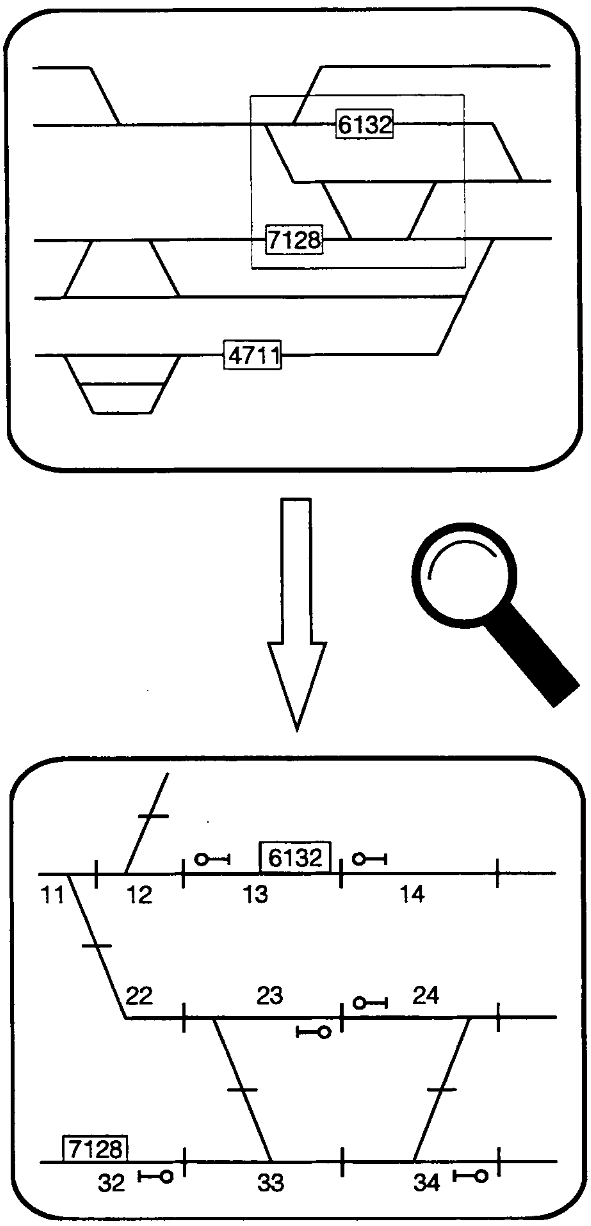


Fig. 2: Nowadays visualisation of the roadway

Today, the train traffic dispatcher is supported in his work by two grafical displays. One, called time-travel-diagram (Hitoshi, 1992; Hundt, *et al.*, 1990) (Fig. 1), shows the time behavior of a train. The other, called network overview (Hundt, *et al.*, 1990) (Fig. 2), shows the actual train position and the roadway reserved for a train. In those days, time-travel-diagrams were plotted by the dispatcher with a pencil on a sheet of paper. Today, this plot is simply done by a computer. The sheet is replaced by a display.

3. INCONVENIENCE OF EXISTING HCIs

The HCIs nowadays used in train traffic control have the following disadvantages (Müller, 1994):

- To perceive conflicts occuring in future, the time-travel-diagram (Fig. 1) must be throughpassed. Reservation conflicts are shown by crossing lines. Those lines first have to be found. Having found a crossing line, the dispatcher needs the network overview (Fig. 2) to verify, if the two trains possibly invoked in a conflict will run on the same roadway. In case of different roadways, the trains will simply pass each other. No conflict will occur. So conflicts yare not so easy to identify.

- To solve a conflict, the dispatcher needs to perceive its reasons. These are difficult to identify in a time-travel diagram.

- As train position (Fig. 2) and time behavior (Fig. 1) are displayed on different displays, the dispatcher has to split his attention. He has to supervise two displays in parallel. Normally he puts his fingers on one display to keep the position of a train and searches its time behavior on the other screen.

- Future is calculated by simple extrapolation of the lines in the time-travel diagram. Often this extrapolation is far away from the real process situation. So the dispatcher do not believe in the preview of the future.

- The importance of the present is neglected. It is only presented by a point in the time-travel diagram.

In case of high train density and high speed, lots of lines will occur on the display scrolling extremely fast. So the dispatcher will be overcharged and will be unable to do a good job. As new traffic systems have a high degree of automation, the dispatcher is no longer needed in case of little conflicts controled by the system itself. In case of big conflicts, the dispatcher will be frightened. His lack of training and the little time resting to solve the problem will stress him. Bad decisions, delays in the transport system and uncontent dispatchers will be the consequences.

4. A NEW DESIGN APPROACH

To find an optimal HCI, a user centered design (Coutaz, 1990; Norman, 1986) is proposed, which takes human mental capacity and his physological abilities (Anderson, 1990; Card, Moran and Newell, 1983) as the most important part of the man-machine-system (Johannsen, 1993; Sheridan, 1992).

To do a good job, the dispatcher must be supported by the HCI as good as possible . This objective can be reached by an optimal job sharing between man and software (Charwat, 1992). While the software is very strong in proceeding big quantities of data and executing programmed problem solutions, the dispatcher is strong to recognize complex correlations and to decide (Rasmussen, 1983). So the big quantity of data has to be visualized by the software in a way that correlations
can be easily recognized by the human visual system (Rock, 1984).

5. DESIGN PROCESS

For an efficient design two important components had to be found out: First the information the dispatcher needs to control trains on the network, second the process values, which form this information. As trains are nowadays controled using the existing HCI, all important information is already presented in it. In discussions with dispatchers and by an analysis of the existing diagrams, process values and information could be recognized. As every process value must be handled by the operation control system, an analysis of the maglev control system (Schnieder, 1983) gave additional values not used in the existing interface. To reduce quantity of information, all values were classified according to their importance for the dispatcher. Using presentation rules proposed by Bertin (1974) and Tufte (1983) they were presented respecting that order in the design. HCI guidelines (Mayhew, 1992; Shneiderman, 1992) were applied to the design of the user dialog. The designs than were improved discussing with dispatchers having experiences in their jobs (van der Schaaf, 1991).

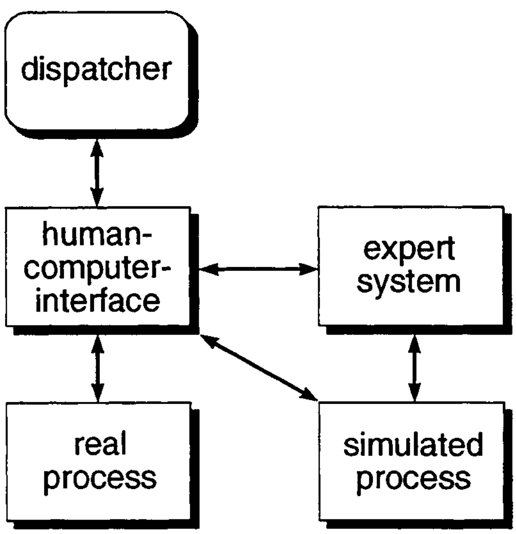


Fig. 3: DISPOS architecture

6. USE OF AN INTELLIGENT ASSISTANT SYSTEM

To make the software supporting the dispatcher the best possible, an intelligent assistant system (Boy, 1991) (Fig. 3) called DISPOS will be developed consisting of an expert sytem and a simulation of the process (Voit, 1994). The expert system finds out different solutions of occuring conflicts in the net and is able to evaluate those solutions. The simulation simulates the traffic on the network based on a model of the operation control system. As this simulation is used the disadvantages of the bad prediction of the future of nowadays interfaces is avoided.

7. PRESENTATION OF NEW DESIGN IDEAS

The proposed new HCI for high speed maglev train control consists of three information layers shown in Fig. 4. The first layer shows a completely new invented diagram to visualize conflicts, its reasons and effects on the network traffic. This diagram will be explained later on. Using this diagram, a dispatcher can easily understand, if there is a critical situation on the network and if he has to interact. Using the time-travel diagrams given on information layer two, he is able to exactly investigate the actual traffic situation. These time travel diagrams are improved compared to those used today:

- Conflicts are marked by a cycle and must no longer be searched by throughpassing the diagram.

- The actual traffic situation is shown directly under the diagram and must no longer be searched in a different diagram.

- The importance of the present situation is respected showing the actual train positions encoded as triangles on the guideway.

- An improved design of the common diagram makes it easier to understand.

The third information layer presents detailed information diagrams the dispatcher needs in special situations like train connection diagrams and velocity diagrams. The shown three layers are used in parallel.

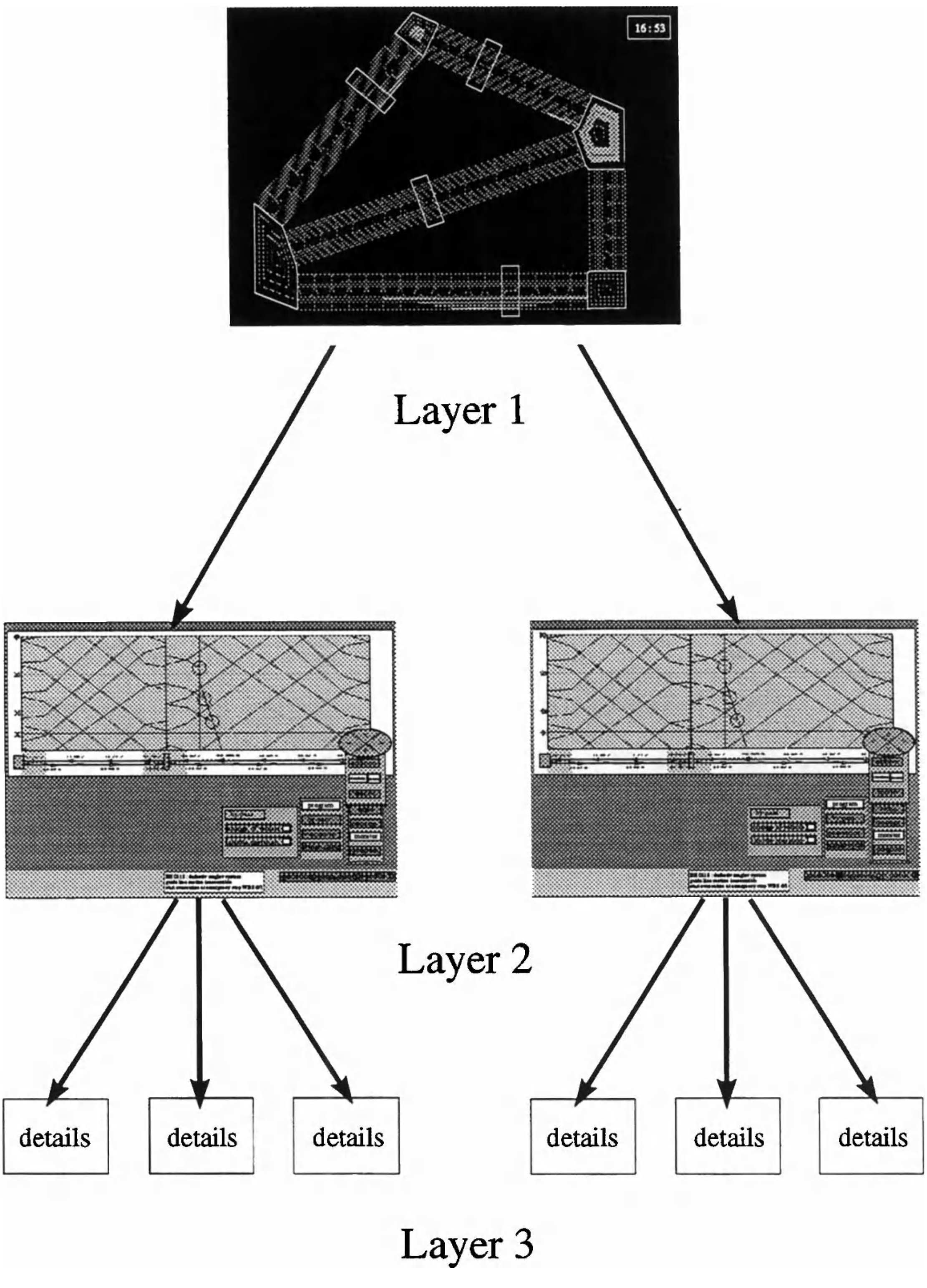


Fig. 4: A three layer information hierachy

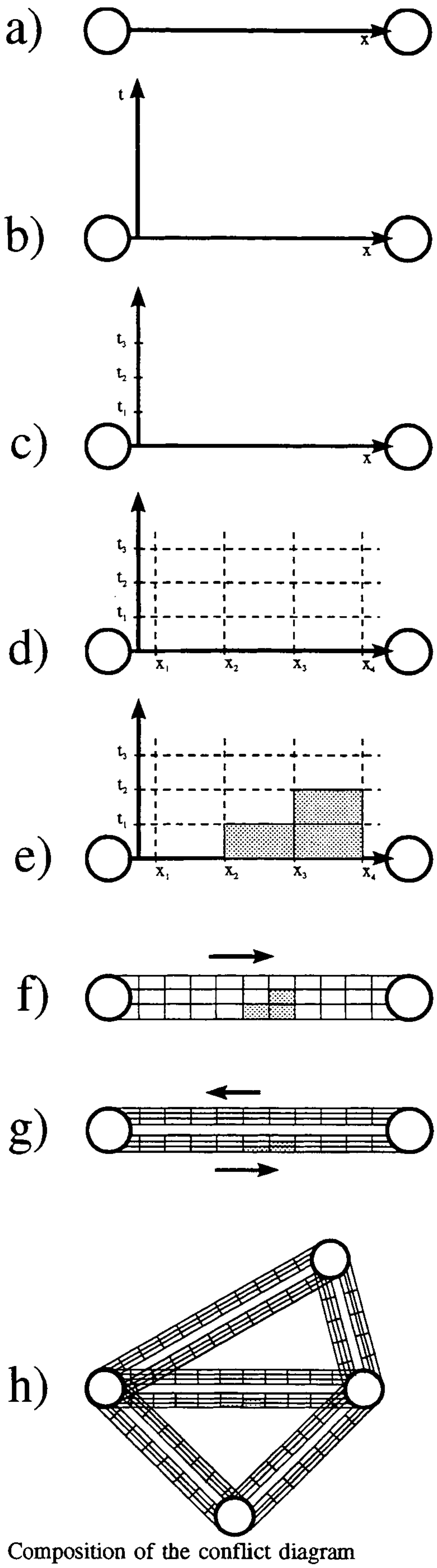


Fig. 5: Composition of the conflict diagram

8. CONFLICT DIAGRAM

To show the development of conflicts on the network and its effects on the traffic situation the following design idea was invented (Fig. 5).

a) Two stations are linked by a line giving the position x of a train.
b) A time axis is introduced to represent present and future.
c) The time axis is quantized.
d) The postion is quantized, too.
e) The trains passing a bloc are counted and are compared to those that should pass this bloc according to the time-table. If a conflict occurs, the difference between actual and nominal number of trains in that bloc is calculated and coded by colors. Using the colors of a fire, e.g. white for big difference, red for small difference, conflicts and their intensity are spreading like a fire over the network.
f) X- and t-axis are integrated in the guideway.
g) Two guideways for two directions are designed.
h) A network of guideways is constructed.

9. DISPATCHERS WORKING AREA

To give the dispatcher a more interesting job the use of his knowledge is forced and simple routine jobs are done by the computer. For he realizes the importance of his tactical decisions the concept of working space must support his importance. So the concept of nowadays computer working areas in process control tasks has to be reviewed. The aspect of team work becomes important. As interactions of the dispatcher are not frequent, he should not be tied any longer to his monitor. It is proposed to visualize the conflict situations on the network on a big screen projection (first information layer) (Fig. 4), which can be perceived by multiple dispatchers in the control tower to discuss about it. Information layer two should be presented using a large screen on a desk individual for each dispatcher. Interacting with this screen by a lightpen he should be able to get information layer three on a portable pad. This pad can be moved around. Doing all interaction on that pad, he can interact without sitting at his desk. Even if a pad like that do not exist today, it could be constructed in near future. To evaluate this working area concept hardware modells have been build.

10. CONCLUSION

It has been shown that HCIs existing nowadays in train traffic control have reached their limits and can not do a good job any longer. As train speed and density increases human mental workload and human mental capacity have to be considered (Anderson, 1990; Card, Moran and Newell, 1983). This leads to completely new display designs to represent important information for the dispatcher can perceive it easily. Using a user centered design approach (Norman, 1986) not only the displays but also the working area has to be redesigned for the dispatcher feels glad to do his job. The ideas described in this paper introduce designs to high speed train control which allow the dispatcher to do his job the best possible.

AKNOWLEDGEMENTS

This research project was sponsered by the German Ministery of Research and technology BMFT-Nr. 514-7291-TV 9405.

REFERENCES

Anderson, J. R. (1990). *Cognitive Psychology and its Implications*. W. H. Freeman & Company,New York.

Bertin, J. (1974). *Grafische Semiologie (Sémiologie Grafique)*. Walter De Gruyter & Co.,Berlin.

Boy, G. (1991). *Intelligent Assistent Systems*. Academic Press, London.

Card, S., T. Moran and A. Newell (1983). *ThePsychology of Human-Computer Inter action*. Lawrence Erlbaum Associates.

Charwat, H. J. (1992). *Lexikon der Mensch- Maschine-Kommunkation*. Oldenbourg Verlag München, Wien.

Coutaz, J. (1990). *Interfaces Hommes-Ordinateur*. Dunod informatique, Bordas, Paris.

Hitoshi and IEDA (1992). Computer-aided Evaluation System of Train Operation Schedule-IEDIS-*,Japanese Railway Engineering*, **No.117**, pp. 6-9, 1992, March,

Hundt, R. *et al.* (1990). Rechnerunterstützte Zugüberwachung in Köln und Hannover, *Signal + Draht*, **7/8**, pp. 141-144.

Johannsen, G. (1993). *Mensch-Maschine-Systeme*. Springer Verlag Berlin Heidelberg.

Mayhew, D. J. (1992). *Principles and guidelines in software user interface design*. Prentice Hall, Englewood Cliffs, New Jersey.

Müller, J.-O. and E. Schnieder (1994). New Display-Designs for Visualisation of Maglev-Transit-Operations, in T.K.S. Murthy, *Computers in Railways* IV - **Vol. 2**, Railway Operations, S.479-486, Computational Mechanics Publications, Southampton, Boston. Norman, D.A. and S.W. Draper (1986). *User centered system design*, Lawrence Erlbaum Associates, Publishers.

Rasmussen, J. (1983). Skills, Rules and Knowledge. Signals, Signs, and Symbols, and other Distinctions in Human Performance Models. *IEEE-SMC* **Vol. 13**, No. 3, pp. 257-266. May/ June, 1983.

Rock, I. (1984). *Perception*. Scientific American Books, Inc., New York.

Schnieder, E.(1983): Requirements of operations control for maglev transit systems, *Proc. of the 4th IFAC/IFIP/IFORS-conferenceon control in transportation systems*, pp. 89-93, Baden-Baden.

Sheridan, T. B. (1992). *Telerobotics, Automation, and Human Supervisory Control*. The MIT Press, London, England.

Shneiderman, B. (1992). *Designing the User Interface*. Addison-Wesley-Publishing Company Inc., New York.

Tufte, E. R. (1983). *The Visual Display of Quantitative Information*. Graphics Press, Box 430, Cheshire, Connecticut 06410.

van der Schaaf T.W. (1991). Einbeziehung von Benutzern bei der Gestaltung graphischer Anzeigen, Internationales Kolloquium. *Einsatz neuer Informations- und Leitsysteme in Verkehr, Prozeßführung und Fertigung*, pp. 153-163, TÜV Reinland, Köln.

Voit, F., J.-O. Müller and E. Schnieder (1994). DISPOS - Ein System zur Dispositionsunterstützung im Magnetschnellbahnverkehr, VDI- Gesellschaft Fahrzeug-. und Verkehrstechnik, *Fachtagung "Aktive Leittechnik für Schienenfahrzeu - ge"*, S. 113-122, 8.9.-9.9.1994, Aachen.

ADVANTAGES OF MASS-DATA-DISPLAYS IN PROCESS S&C

C. Beuthel, B. Boussoffara, P. Elzer
(carsten, badi, elzer@ipp.tu-clausthal.de)
Institute for Process and Production Control Techniques of the Technical University of Clausthal, Leibnizstrasse 28, D-38678 Clausthal-Zellerfeld, F.R.G.

K. Zinser, A. Tißen
(zinser, tissen@decrc.abb.de)
ABB, ASEA BROWN BOVERI, Corporate Research, D-69123 Heidelberg, F.R.G.

Abstract: The paper presents the adaptation of a method for presenting masses of scientific data to process S&C. Experimental evidence shows that this type of Mass-Data-Display (MDD) is superior to P&I diagrams with respect to the detection time of abnormal process situations. The respective experiments are described in the paper.

Keywords: Displays, Man/machine interfaces, Process control, Supervisory control

1. INTRODUCTION

The improvement of the quality of human decision making under critical conditions has been one of the foremost goals of research in man-machine-systems during the past decades. Several techniques for achieving this goal have been developed and successfully tried. One of these techniques is the use of advisory expert systems as e.g. implemented in the GRADIENT system (Elzer, et al., 1989). Besides direct advice, given to the operator, such systems also fulfil the purpose of "buying time" (Elzer, et al., 1992). This reduces the danger of the operator panicking because of time pressure and thus taking inadequate action.

But, like all techniques that rely on predefined proposals for actions, expert systems can only give advice in situations that have been foreseen by the designers of the industrial process or the S&C system. Therefore they are of little help in unknown situations.

2. THE "MASS-DATA-DISPLAY" (MDD)

2.1 The Principle

The authors have therefore tried to identify a technique that enables the operator to maintain an overview over the entire process with moderate effort and thus to develop a "feeling" for the overall behaviour of the process and upcoming abnormal situations. In contrast to other approaches that try to compress all information about the process into very few symbols (like e.g. the "Chernoff Faces" (Chernoff, 1973) it has been tried to display **all** process values in a form that allows to utilize the pattern matching capabilites of humans to their advantage. The approach chosen is mainly based upon the work of Grinstein (Grinstein and Bergeron, 1989).

The underlying principle is that each process value is represented by a small icon on a display screen. This icon changes its shape (size, colour etc) according to the changes of the process value. The operator does not observe the behaviour of each individual icon, but the overall pattern that develops because a number of icons - that belong to technologically interrelated process entities - change their appearance simultaneously. This technique allows to show an overview over several hundreds (even thousands) of values on one display screen of average size. Several models of such displays were built in cooperation between IPP-TUC and ABB research and one version exhibited at the "Interkama" technical trade fair in Duesseldorf in 1992 (Zinser, 1993). Fig. 1 shows such a mass-data-display of a coal-fired power plant.

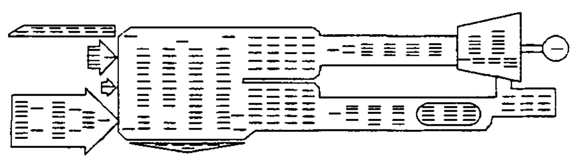


Fig. 1. A Mass-Data-Display of a coal fired power plant

The results so far have been very encouraging. Experiments have shown that this technique allows the identification of abnormal process states by an order of magnitude faster than traditional display techniques and that the acceptance by test persons is very good.

2.2 Implementation of MDD's

Design problems. But the design of this type of mass-data-display also poses many interesting problems with respect to the details of its implementation. Grinstein's original proposal of the "five limbed stick" was somewhat difficult to use with typical process values. It basically requires that the values represented by the "limbs" are closely interrelated, like e.g. velocity, pressure, temperature, mass transfer and density at a particular point in a flow field.

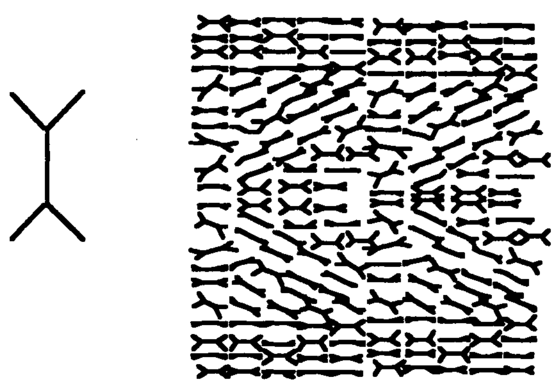


Fig. 2. Grinstein's "five limbed stick" (from: Grinstein and Bergeron, 1989)

But this can not be guaranteed by the S&C-system of a typical technical process. For reasons of cost efficiency only those values are measured and made available to the operator that are regarded as essential for the proper supervision of the process. Therefore each available process value is to be regarded as important in its own right and presented individually.

Besides, calculations soon showed that technical processes very rarely provide enough values to produce a real "pattern" on a display screen, if they are not shown individually.

The appearance of the individual icon. Therefore the discussion concentrated on shape, colour, or brightness of individual icons, connected to one process value each. This discussion soon included the "sympathy value" of an icon with respect to shape, brightness or colour. The rationale was that a pattern should very soon "look awkward" after some values left their normal range in order to warn the operator in a conspicuous manner. Fig. 3 gives an impression of the many possibilities that were discussed.

The correlation of the parameters of the icon with the process values. This aspect was also very intensely discussed. It turned out that only experiments could decide, which was the best solution. So, e.g. it is not always quite easy to decide, whether one or more parameters of an icon should be connected to a process value and thus change according to changes of that value.

A particularily difficult item is the "mapping function" between the time behaviour of a process value and the related parameters of an icon. It turned out that deviations from the normal state of a process can be discovered most rapidly and reliably when the icons are "normalized", i.e. have a uniform appearance in normal process states. Fig 4 shows the effect. The individual icons are short lines that are normalized horizontally. Their breadth and the angle towards the horizontal axis are proportional to the deviation of the related process value from its normal state.

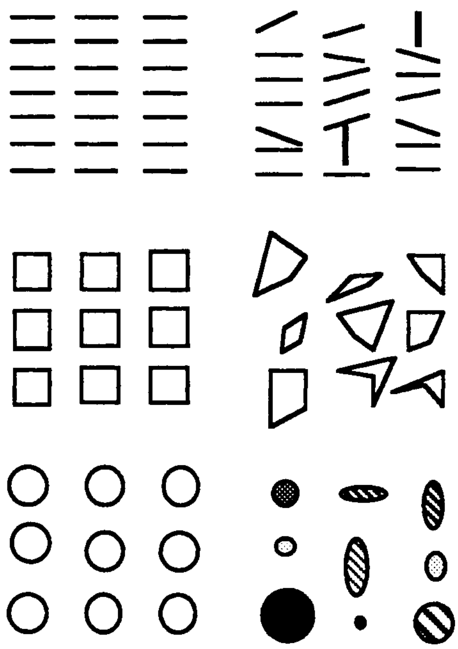


Fig. 3. Some examples for elements of MDD's

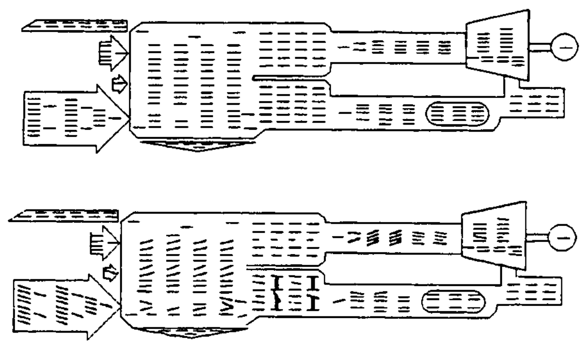


Fig. 4. Example of a "normalized" MMD in a normal and an abnormal process state

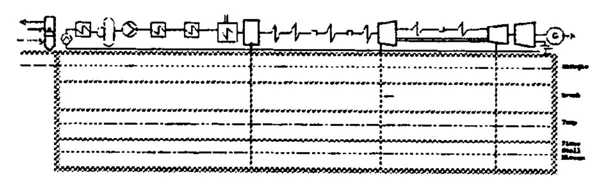


Fig. 5. A table presentation of the process

The layout of the entire overview picture. The order and grouping of a large number of icons into an overview picture of the process has also to be determined rather carefully. Various criteria can be applied which lead to a completely different layout of the display:

- topological relationship
- flow of material, energy or information
- aesthetic considerations

The example in Figs.1 and 4 shows a display layout following a "topological order" of the icons. It is based on a simplified representation of the water steam cycle of the power plant. The icons are positioned according to the approximate places of the respective sensors in the real process.

Fig. 5 shows a completely different approach. The water steam cycle is presented in a more symbolic (table) form. The process values are ordered in four horizontal rows: energy, pressure, temperature, and setpoint levels. This representation was favoured very much by experts, but the experiments showed that it caused some difficulties for less trained persons.

3. EXPERIMENTS

3.1 General

From the beginning it had been clear to the authors that the discussion of the expected advantages of this rather novel display form had to be supported by experiments and quantitative measurements.

Therefore two types of experiments were planned and performed:

- One, comparing the performance of operators using such mass-data-displays with those using "classic" P&I diagrams, and
- a second one, comparing the performance of operators using various different versions of MDD's.

The first set of experiments was perfomed with process experts as test persons and also included in-depth discussions of the complementary roles of the new method and the "conventional" forms.

3.2 The Experimental Setup

The experiments were jointly planned by ABB and IPP-TUC and conducted at IPP-TUC.

The experimental setup consisted of a very elaborate computer simulation of a coal-fired power station that was available at ABB, a set of P&I diagrams designed for controlling the simulated power station, and implementations of five types of MDD's.

These MDD's were selected according to the following considerations:

As a major item of the design discussion had been whether the dynamic behaviour of the icons should be related to normalized or non-normalized process values, at least one of the displays had to be non-normalized.

Among the normalized versions there had to be at least one linearized (table) version.

The remaining three displays - all normalized with topologically ordered icons - should differ as much as possible in the appearance of the icons. Therefore it was decided to use:

- short lines, changing width and direction,
- circles, changing diameter and brightness,
- adjacent squares, changing brightness.

Table 1 gives an overview and shows the correlation with the display numbers.

Table 1 Types of MDD's investigated

Nr.	layout	icons	mapping func.
#1	topological	lines	normalized
#2	table	lines	normalized
#3	topological	squares	normalized
#4	topological	squares	non normalized
#5	topological	circles	normalized

Fig. 6 explains the overall systematics of the selection of the various types of displays for the experiments. The numbers in the circles represent the numbers of displays tested of this kind.

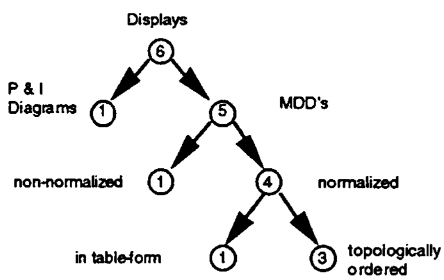


Fig. 6. Systematic selection of display types

Three abnormal situations were presented to the test persons and the time for the correct identification of these situations measured:

1 Leakage in a high-pressure preheater,
2 Valve spindle broken at a high-pressure preheater,
3 Feedwater pump trip during increase of power output.

As a cross-check a fourth situation was presented to the test persons: a load transition of +/- 10 %. This was chosen because it produced changes of the display pattern that were similar to that of abnormal status.

The test persons were students of the TUC. There were three groups of four students with different levels of knowledge about technical processes:

1 Beginners (1. grade students)
2 Persons with basic knowledge (B.A. completed)
3 Well trained persons (working at master's thesis)

All groups were given an individual group training of two days, in which the general characteristics of the power station process, possible abnormal cases and the use of the interface were explained. The test persons then had the opportunity to get acquainted with the setup for three hours.

The experiments consisted of five sessions of two hours each, in which the abnormal cases were presented to the test persons in randomized order and the time for the correct identification of these situations was measured. In parallel, the behaviour of the test persons and their comments during the experiments were recorded on videotape. 11.000 individual measurements were collected. Afterwards the results were evaluated using a commercial software package for statistical analyses (SAS©).

Two different methods were used to analyse the test data: parametrical tests for normally distributed data (analyses of variance) and non-parametrical tests for not normally distributed data. Because it can not be presupposed that the data are normally distributed, the non parametrical test were used to verify the parametrical tests. The non parametrical tests were conducted with box-plots (Johannsen, 1993, Schlotzhauer, 1987, Montgomery, 1991, Sachs, 1992).

3.3 Some Results

It turned out that all test persons identified the abnormal situations much earlier when using the MDD in contrast to the P&I diagrams. In one case they were even 20 times faster. This result clearly shows that the design purpose of the new method, i.e. to "buy time" for the operators in order to reduce stress and to enable them to make qualitatively better decisions, has been clearly achieved. The results even exceeded the expectations of the authors. Fig. 7 shows a simplified plot of the results under situations 1 (leakage in a high-pressure preheater) and 2 (Valve spindle broken).

But there were also some interesting differences in the performance of the test persons that were correlated to differences in the design of the various MDD's. Due to the limitation of the length of this paper only a few can be presented here.

Among these very interesting insights the clearest difference between the various types of MDD's could be observed between the topologically organized displays on one hand and the table form on the other. For one group of test persons, the beginners, the table display even caused reaction times 50% longer than the other types.

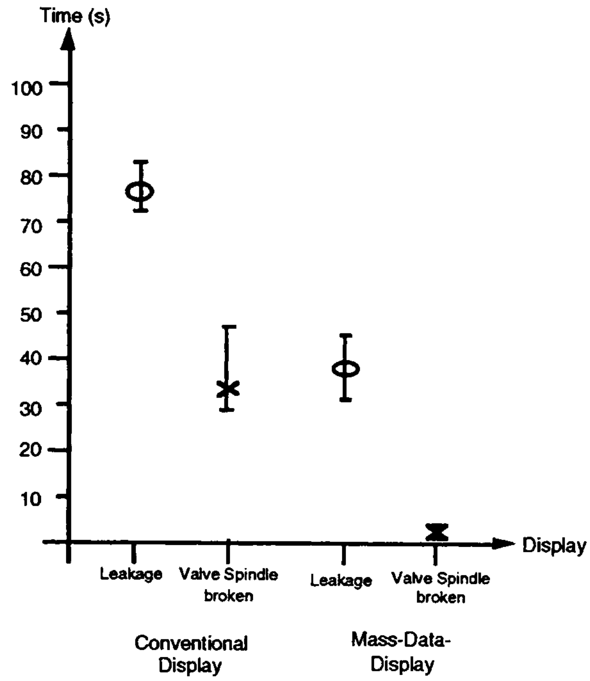


Fig. 7. Comparison of the recognition time for abnormal situations using MDD's and P&I-diagrams

Of course, compared to P&I diagrams, the measured values are still extremely good. The reaction times decreased with knowledge level of the test subjects. Fig. 8 shows the respective plot.

Another type of MDD showed inconsistent results: the non-normalized display with squares as icons that changed their brightness according to the process value. It was faster than the others in situation 1, but much slower in situation 2. This behaviour was consistent over all three groups of test persons. Fig. 9 shows the results for the "expert group".

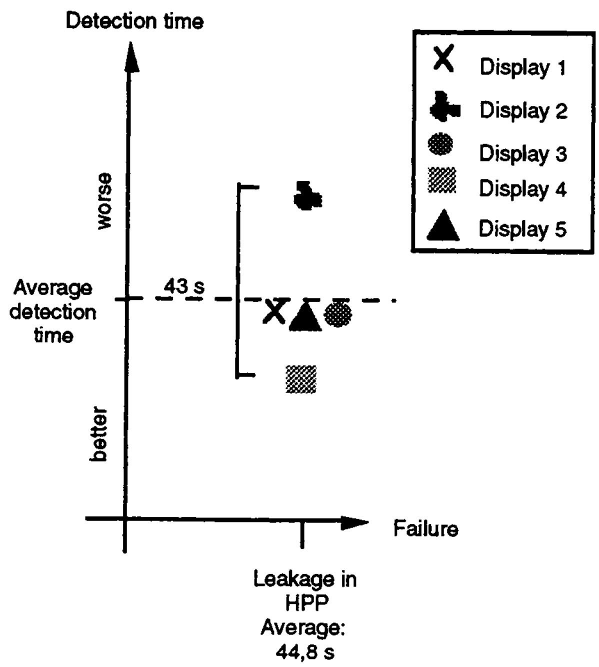


Fig. 8. Insufficient performance of beginners with the linearized display

A closer analysis revealed that the reason was rather subtle: a combination of the "restless" appearance of the display as a whole, that did not allow the test persons to become acquainted with a "normal" appearance, with a non-linear calibration of the mapping of the brightness of the icon onto the process value. Therefore this type of display was ruled out as technically too difficult to calibrate.

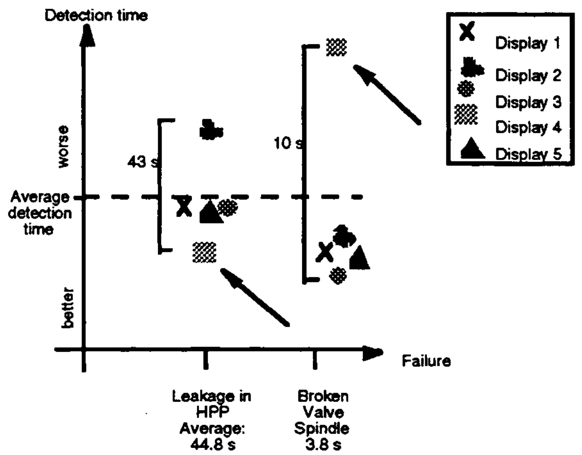


Fig. 9. Inconsistent performance of the non-normalized display

The remaining three types of MDD showed very similar performance. They were all topologically arranged and normalized. Fig. 10 shows a plot of the respective measurements.

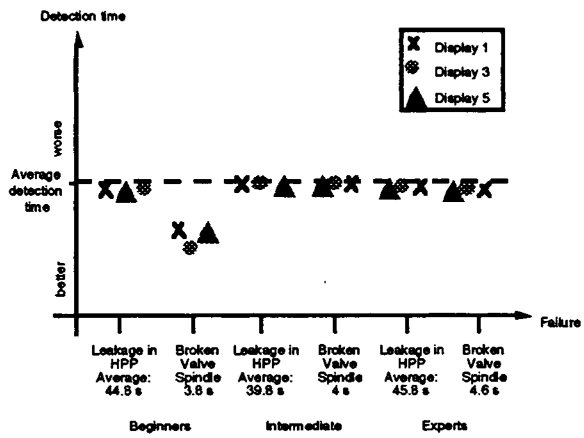


Fig. 10. Comparable performance of topologically arranged and normalized displays

However, they were very differently accepted by the test persons. The display with the gray squares and that with circles as icons caused various irritations. The small lines were generally judged as giving the clearest impression of the behaviour of the process.

There was only a slight problem insofar as the angle of the lines towards the horizontal direction could be mistaken as a tendency instead of a value. But as all test persons got accustomed to the appearance after very short time, this type of MDD was finally selected for further development.

User errors. Another important factor beside the reaction time were the user errors made by the test persons. One had to distinguish between two different types of errors:

a) a failure was not or wrongly identified,
b) a normal process state was mistaken for a failure.

In the first case the user loses time in repairing the failure or does not even identify it, in the second case it triggers the user to look for a disturbance while there is none. This is not that critical if no real disturbance occurs during his search but the acceptance of a display is likely to decrease if this happens too often.

Fig. 11 shows the user errors during the tests. It is conspicuous that most of the errors were made by the beginners group and that overall most of the errors were made by testing the non-normalized display. The table display performed well with experts but beginners and intermediates had problems with it. The three normalized displays under test performed about the same.

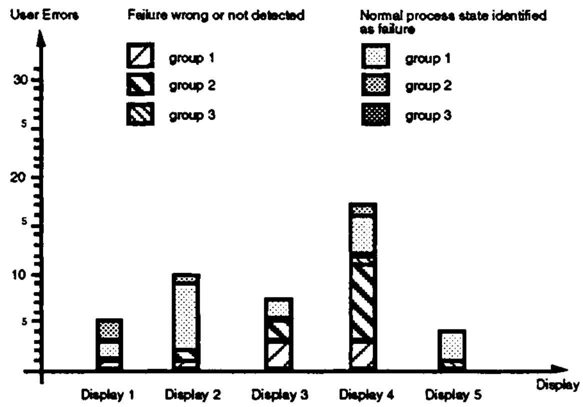


Fig. 11. Number of user errors for various types of MDD

The overall performance of all MDD's was still sufficient. Out of a possible total of 240 errors, fig. 12 shows the actual errors as (%).

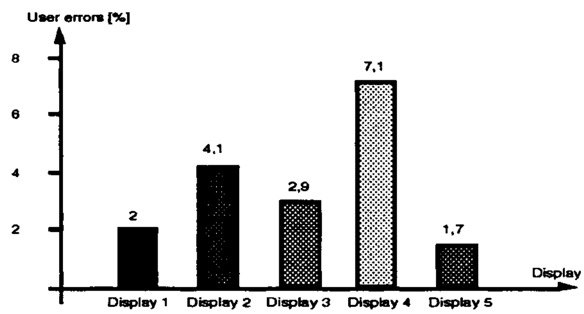


Fig. 12. Percentage of user errors

4. INTEGRATION OF MDD'S IN CONTROL ROOM DESIGN

The discussion showed that the new method should not be regarded as a replacement for the classical P&I diagrams. In order to be able to quantitatively evaluate a situation and to make final decisions the test persons had to use the P&I diagrams, read the individual values, judge the severity of the disturbance and plan the concrete action.

The authors therefore hold that the novel form display should be integrated into a system of process displays in the following way:

- An overview display on the basis of the new mass-data-display method should be visible to the operators all the time, together with the mandatory "event list" in textual form.

- The operators should be able to indicate on this screen the area(s) of the process in which they suspect abnormal behaviour according to the local development of the pattern observed. The MMI should then provide on request selected P&I diagrams of the related areas for closer inspection or allow to invoke an expert system for a possible prediction of the development of the process state (Elzer, et al., 1992).

- The operator should be able to invoke a modified presentation of the mass-display that contains trend indicators showing changes of the process values over time. This will allow the operator to judge the urgency of compensatory actions according to the speed and tendency of changes of the process state.

- In parallel, expert systems (Elzer, et al., 1989) should continuously monitor the process in order to provide the operator with warnings concerning foreseeable abnormal cases.

5. CONCLUSION

It has been shown that mass-data-displays, that were originally designed for the presentation of scientific data, can also be used with advantage for the supervision of industrial processes. Experiments, conducted by the authors, have shown that they help to considerably reduce the time an operator needs to recognize an abnormal situation. Thereby they contribute to a further reduction of operator stress and an increase of overall plant safety. This is particularly true when MDD's are used in connection with "classic" P&I diagrams and knowledge based operator support systems.

The authors therefore hold that MDD's are an extremely useful contribution to the design of human-computer-interfaces for process S&C and will continue experimental research into this subject. This work will include further refinement of the design of MDD's, using the insights that could already be gained from the experiments described in this paper.

REFERENCES

Chernoff, H. (1973). The use of faces to represent points in k-dimensional space graphically. In: *Journal of the American Statistical Association (68), pp.* 361-368.

Elzer, P.F., H. Siebert and K. Zinser, (1989).New possibilities for the representation of process information in industrial control. In: *3rd IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Man-Machine Systems*, June 1988,Oulu, Finland, Pergamon Press, Oxford-New York, 1988, pp. 139-143

Elzer P. F., Zinser, K., and Weisang, C. (1992). Alarm filtering vs. failure prediction. In: *5th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Man-Machine Systems (Preprints)*, Session 3.1.5, June 1992, The Hague, The Netherlands, Pergamon Press, Oxford-New York, Paper 3.1.5.

Grinstein, G. and Bergeron, D. (1989). The visualisation of scientific data. In: *Springer, Informatik Fachberichte 222 (19.GI-Jahrestagung)*,pp.1-9.

Johannsen, G (1993). *Mensch-Maschine-Systeme.* Springer-Lehrbuch, Springer-Verlag Berlin Heidelberg 1993.

Montgomery, Douglas C. (1991). *Design and Analysis of Experiments.* John Wiley&Sons, New York 1991.

Sachs, Lothar (1992). *Angewandte Statistik.* Springer Verlag Berlin Heidelberg 1993.

SAS ©. SAS Institute Inc. , Cary , NC 27513

Schlotzhauer, S. and R. C. Littell (1991). *SAS System for Elementary Statistical Analysis.* SAS Institute Inc. , Cary, NC 27513

Zinser, K. (1990). Moderne Prozeßleitwarten: Neue Aspekte der Gestaltung, Ausstattung und Informationsdarstellung. In: *atp- Automatisierungstechnische Praxis (32) 12,* pp. 579-585.

Zinser, K. (1993). Neue Formen und Medien der Prozeßvisualisierung. In: *atp- Automatisierungstechnische Praxis (35) 9, pp.* 499-504.

Zinser, K. (1994). Multimedia's push into power. In: *IEEE-Spectrum, The Moon Landing, A 25th Anniversary Perspective, July, pp.* 44-48.

A GENERIC TASK FRAMEWORK FOR INTERFACE ANALYSIS AND DESIGN IN PROCESS CONTROL

Jean-Yves Fiset

Centre de recherche informatique de Montréal
1801 ave McGill College, Bureau 800, Montréal (Qc) Canada H3A 2N4

Abstract: The analysis and design of human-machine interfaces in process control applications often fail to take into account the information requirements associated with specific categories of tasks. This paper defines generic tasks (GTs) which are useful for describing the interaction between an operator and a complex system, and to analyze and specify the information content of the interface. The GTs--detection, transition, tuning, compensation, and diagnosis--and their syntax are expressed using MAD (Méthode Analytique de Description de tâches), a hierarchical, object-oriented, formalism.

Keywords: process control, human factors, interfaces

1. INTRODUCTION

Empirical results found in the literature show that operators use different types of information, depending on the kind of task they carry out (Kieras, 1992; Landerweerd, 1979; Vermeulen, 1987). Yet, the analysis and design of human-machine interfaces for process control applications frequently fail to take into account the kind of task that is being carried out. This is also the case for existing methods for interface analysis or design such as ecological interface design (EID) (Vicente and Rasmussen, 1992) or the operator function model (OFM) (Mitchell and Miller, 1986). Further, while research results have been reported comparing the relative merits of graphical, analog, and digital data display for process control, vague or inconsistent definitions have often been used to characterize the tasks for which the results were obtained. This makes the interpretation and integration of these results difficult, thus limiting their potential usefulness in the analysis and design of interfaces.

This paper defines generic tasks (GTs) to describe more precisely a significant portion of the interaction between an operator and a complex system. The Gts and their syntax are expressed in MAD (Méthode Analytique de Description de tâches), a hierarchical, object-oriented task description formalism. Integrating the GTs with MAD provides a formal framework to describe an operator's tasks and information requirements. The GT framework supplies vital information for analyzing or designing the interface, and its syntax helps analysts and designers in checking for the completeness of the interface. The GT framework also facilitates the interpretation and integration of existing and future empirical research results.

2. DEFINITION OF THE GTs

A set of tasks suitable for describing a large repertoire of process control situations has been identified after a survey of the published literature. These tasks (i.e., the actual GTs) were adopted since they had been partly validated in modeling human problem solving when controlling a complex system (Knaeuper, 1983). The GTs are:

- Detecting: detecting the occurrence (or lack of) of an event.
- Transitioning: taking a system from one state to another (e.g., start-up, production shift). Normally accompanied by a simultaneous detection task.

- Tuning: making small corrections to a system to keep it within an operating region. Normally accompanied by a simultaneous detection task.
- Diagnosing: determining the cause of an event.
- Compensating: making changes in response to a failure or to an emergency.

The GT framework also integrates relevant aspects of the human performance model originally used with the tasks (Rouse, 1983). The model specifies successive categorization, planning, and execution steps for each task, and shows how a task can be carried out through 'standard' or 'non-standard' means. This makes a description based on GTs especially robust. For example, carrying out a transition GT involves:
- Categorizing (recognizing) the transition GT to be carried out through pattern-matching if possible, or else through more elaborate reasoning.
- Planning the execution of the transition GT, through re-use or adaptation of an existing plan, or by developing a new plan.
- Executing and supervising the execution of the plan, either through usual means or ad hoc means.

The model enables a finer grained description of each GT and allows its information requirements to be specified. For a transition GT, the planning step is most crucial. In 'standard' situations, the content of a plan or procedure to carry out the transition GT can be specified by leveraging relevant empirical knowledge. In 'non-standard' situations (e.g., equipment failure), knowledge can be specified to allow an operator to elaborate a plan. For transition GT, this requires, among other, knowledge about the causality and dynamics of the system variables which can affect the GT's goal. Specifying these information requirements has been discussed more completely elsewhere (Fiset, 1993; Fiset, 1994).

The next section introduces the MAD formalism.

3. THE MAD FORMALISM

MAD prescribes how to express a task as an object or set of objects (Pierret-Golbreich et al., 1989). MAD descriptions have been found to be practical, systematic, and to lead to a better completeness in the task description in comparable applications (e.g., air traffic control) (Sebillotte and Scapin, 1992). The main components of a MAD object are a task descriptor, the initial and final states of the world (before and after the execution of the task), the task's goal, the preconditions to trigger the task, the postconditions after the task execution. the sub-tasks and activities that make-up the task, and the plan describing how the sub-tasks are articulated. The plan describing the coordination of the tasks is called a constructor. An example constructor is SEQ (sub-tasks are carried out sequentially); additional constructors can be defined as needed. Extensions needed to fully express the GTs in MAD will be introduced in a later section.

4. EXPRESSING THE GTs' SYNTAX WITH MAD

In this paper, the emphasis will be on expressing the syntactic aspects of the GTs. This syntax is extremely useful for checking the correctness and completeness of the analysis or design of a given interface. The syntax is identified by describing and justifying the relevant properties of the GTs, and by showing the resultant MAD representations. The constructors of interest will be PAR (the GT's sub-tasks can be executed in parallel or simultaneously), SEQ (sequentially), and COND (conditional).

GTs collective and individual properties are identified by considering the purpose of each GT, and their logical and temporal relationships. This is done in the following sections.

4.1 *GTs collective properties*

The original work from which the GTs are drawn did not limit the system size to which these tasks applied (Knaeuper, 1983). This leads to:
- Each GT can be applied to a system, a sub-system, etc., as appropriate, i.e., GTs apply at any level in the control or supervision of a given system.

As mentioned earlier, the information requirements specified for each GT's must support the elaboration of alternate plans to carry it out, which yields:
- Each GT can be carried out through 'standard' means, or through 'non-standard' means.

These collective properties will be reflected in the syntax that is defined in the next section.

4.2 *GTs individual properties and syntax*

Detection GTs detect that an event did or did not occur. In the GT framework, their scope is limited to detecting events affecting the achievement of the goals for the tasks to which they are associated. Detection GTs are thus normally active while preparing or carrying out transition or tuning GTs. By definition, they also precede compensation or diagnosis GT (one cannot compensate or diagnose a fault that has not been detected). Finally, a detection GTs can be carried out by carrying out lower level detection GTs (e.g., at sub-systems levels). These considerations yield:
- Detection GTs are active at the same time as tuning or transition GTs. Figure 1 shows the resultant syntax.

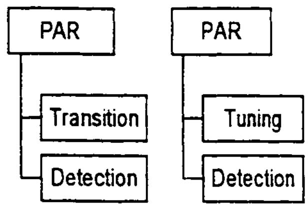


Figure 1: Syntax for detection, transition, and tuning GTs.

- Detection GTs precede either compensation or diagnosis GTs (Figure 2).

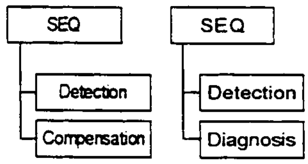


Figure 2: Syntax for detection, compensation, and diagnosis GTs.

- Detection GTs can be broken down into lower level detection GTs (Figure 3).

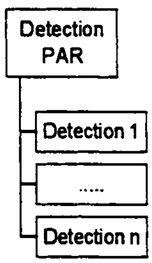


Figure 3: Syntax for decomposing detection GTs.

Transition GTs take a system from a starting state to a target state. Here, state implies an operating region (e.g., 'operating at 100 % of rated power'), not a control engineering meaning. Transition GTs are usually well-defined for a system. Transitions on a system usually imply transitions on its sub-systems. Also, after a sub-system or system has been transitioned, it often has to be monitored and 'tuned' for correcting minor discrepancies. This yield:
- Transition GTs are usually goal driven, i.e., proceed according to a plan.
- Transition GTs can be broken down into lower level transition and tuning GTs (Figure 4; the associated detection GTs have been omitted for clarity).

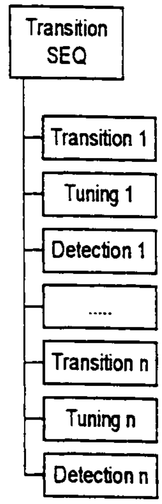


Figure 4: Syntax for decomposing transition GTs.

Tuning GTs are usually 'skill-based' in that they are triggered from cues provided by the environment about minor system deviations. Applying a tuning GT to a system may imply applying tuning GT(s) to some sub-system(s). The resultant properties are:
- Tuning GTs can be broken down into lower level tuning GTs (Figure 5; the accompanying detection GTs have been omitted for clarity).

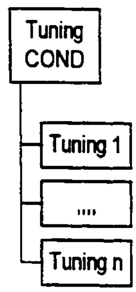


Figure 5: Syntax for decomposing tuning GTs.

Diagnosis GTs are used to determine the cause of an event. Depending on the purpose of the diagnosis (e.g., identifying the equipment to repair, taking compensatory action), it may be carried out to varying levels of detail. Further, diagnosing often implies carrying out test sequences, possibly even carrying out transition GTs on sub-systems. Depending upon the organization of the information, and the depth and purpose of the diagnosis, diagnosis GT can be seen as a recursive application of a detection GT down a detection GTs tree.

The diagnosis GT is a special case in the GT framework. Because of the variety of strategies available, several properties and syntax can be derived. The choice of the appropriate representation for a given analysis must be guided by a clear understanding of the purpose of the diagnosis. The properties are:
- Diagnosis GT can be broken down into lower level GTs (Figure 6).

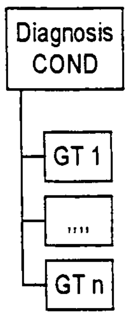


Figure 6: Syntax for decomposing diagnosis GTs.

- In a hierarchy of detection GTs, a diagnosis GT can be seen as a recursive application of a detection GT to each level of the hierarchy (Figure 7).

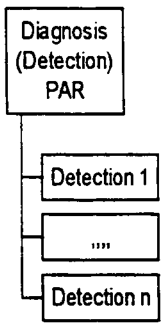


Figure 7: Syntax for decomposing diagnosis GTs into detection GTs.

- Diagnosis GT may be pursued as far as needed for identifying an appropriate compensation GT for dealing with the failure (Figure 8).

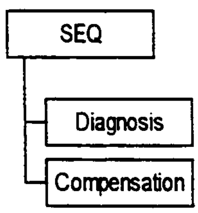


Figure 8: Syntax for diagnosis and compensation GTs.

Compensation GTs are pre-made plans for dealing with a failure. As such, they normally imply derating equipment, changing operating setpoints or configurations, etc. Because of the need to make changes to take care of the failure, compensation GTs normally involve carrying out transition GTs. This yield the following properties:

- One or more detection or diagnosis GT(s) always precede a compensation GT (shown in Figure 8).
- A compensation GT can be broken down into transition GTs (Figure 9; the accompanying detection GTs have been omitted for clarity).

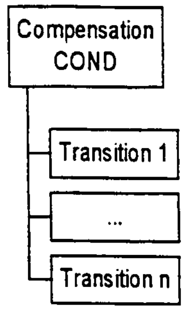


Figure 9: Syntax for decomposing compensation GTs.

The GTs and their properties are still being refined. However, they can already be used to represent an operator's interaction with a system. An example is shown in the next section.

5. EXAMPLE APPLICATION

The system consists of reservoirs A and B, each containing raw materials that are combined into reservoir C. These three sub-systems (A, B, and C) constitute the overall system (Figure 10).

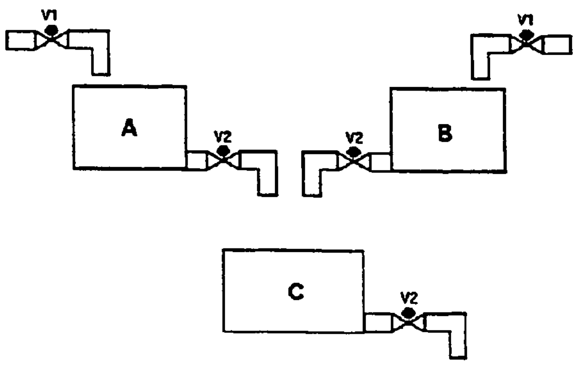


Figure 10: Production system for example.

This system enables the production of a mixture of a given ratio (raw material A/raw material B), at a given thruput (l/min). In this example, the system is to be started-up with all reservoirs empty, and to be brought up to a total thruput of 500 l/min with a balanced mixture of the raw materials (50 % of each). The system goals to achieve are: get a thruput of 500 l/min for V2 (C), get a thruput of 250 l/min for sub-system A, and get a thruput of 250 l/min for sub-system B. Each of these goals can be further decomposed into sub-goals, sub-sub goals, etc., to each of which a GT can be associated. Each GT may specify two different information contents:

- The information needed by the operator to achieve the goal(s) for 'standard situations', i.e., which GTs to carry out, which plan to use.
- The information needed to devise alternate plans to achieve the goal(s) for 'non-standard' situations, should the standard plan no longer apply (e.g., because of an equipment failure).

The latter requires to extend the MAD objects by providing a 'pointer' to a body of knowledge used to devise plans to achieve the goal(s). This knowledge depends on the goal(s) and on the GT(s) being carried out. It is not always required or possible to develop this type of information content for a GT (e.g., if a goal is optional). The GT description for starting up subsystem A is shown in Figure 11.

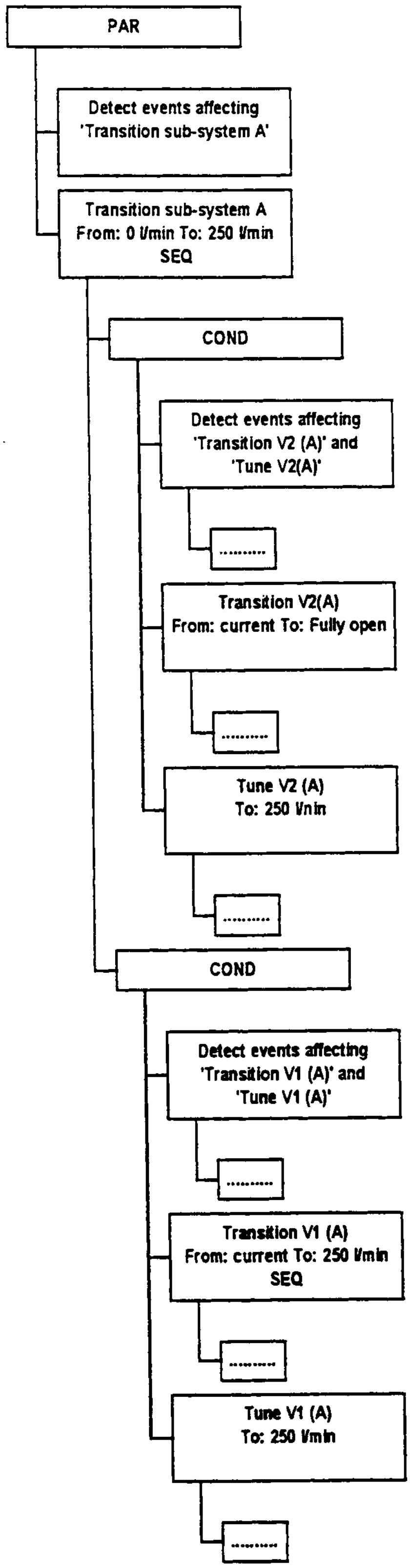


Figure 11: Description of start-up for subsystem A.

Figure 11 shows a rigorous description of the task, achieved through the use of the GTs and syntax described earlier.

Each of the GTs or MAD objects shown (the large boxes in the Figure) contain the information required to achieve its goal(s) according to the standard plan. This standard plan is represented by the GT decomposition and the associated constructor.

When a GT is related to a critical goal, it contains information (or a pointer to the information) on how to devise alternate plans to achieve that goal. This may imply that sub-GTs (and sub-goals) will be abandoned as novel ways to achieve the critical goal will be developed.

This representation also provides a promising formal framework to help dealing with a problem that has plagued human-machine interfaces in process control applications, i.e., alarm flooding. In traditional human-machine interfaces, alarms are non-contextual, i.e., that are triggered when variables cross thresholds (whether fixed or dynamic). Thus, transients or even normal and well-behaved situations like shutdown or start-up often lead to a large number of alarms being generated. The syntax for detection GTs shows that lower level detection sub-GTs can be subsumed by a higher level, goal-related, detection GT. This holds the potential to reduce the alarm flood brought about by goal-irrelevant alarm signals.

6. CONCLUSION

This paper has proposed generic tasks as a solution to the lack of consideration given to task-specific information requirements in the analysis and design of human-machine interfaces in process control applications. The GTs--detection, transition, tuning, compensation, and diagnosis--have been drawn and adapted from the published literature. GTs specify which information must be provided to an operator to achieve goals during normal situations by using a standard plan. When abnormal situations arise during which standard plans are inappropriate and critical goals must be achieved, GTs provide references to knowledge to enable alternate plans to be developed.

For each GT, properties and associated syntactic structure have been identified. The GTs and their syntax have been represented with MAD, a hierarchical, object-oriented formalism.

An example involving the start-up of a hypothetical system has shown how the GTs can be used to formally describe the interaction between an operator and a system.

GTs also provide a framework to help dealing with the alarm flooding problem often found in human-machine interfaces in process control applications.

7. ACKNOWLEDGEMENTS

The work presented in this paper benefited from discussions with former colleagues from the Human-Machine Systems Development Section, at AECL Research, Chalk River Laboratories, Ontario, Canada.

8. REFERENCES

Fiset, J.-Y. (1993). Display design for process control: integration of empirical results. *Procs. 26th Annual Conference of the Human Factors Association of Canada*, 37-42

Fiset, J.-Y. (1994). Generic Tasks for Display Design in Process Control. *Procs. IEA '94, 334-336*

Kieras, D. (1992). Diagrammatic displays for engineered systems: effects on human performance in interacting with malfunctioning systems. *Int. J. Man-Machine Studies*, **36**, 861-895

Knaeuper, A.E. (1983). *A model of human problem solving in dynamic environments*. Report no. 83-3, center for man-machine systems research, Georgia Institute of Technology

Landerweerd, J.A.(1979). Internal representation of a process, fault diagnosis and fault correction. *Ergonomics*, **22(12)**, 1343-1351

Mitchell, C.M. and R.A. Miller (1986). A Discrete Control Model of Operator Function: A Methodology for Information Display Design. *IEEE Transactions on Systems, Man, and Cybernetics*, **SMC-16(3)**, 343-357

Rouse, W.B. (1983). Models of Human Problem Solving: Detection, Diagnosis, and Compensation for System Failures. *Automatica*, **19(6)**, 613-625

Pierret-Golbreich, C., I. Delouis and D.L. Scapin (1989). Un Outil d'Acquisition et de Représentation des Tâches Orienté-Objet. Rapport de recherche no. 1063, INRIA.

Sebillotte, S. and D.L. Scapin (1992). From users' task knowledge to high level interface specification. *Procs. WWDU'92, Berlin, 1-4 Sept.*

Vermeulen, J. (1987). Effects of Functionally or Topographically presented Process Schemes on Operator Performance. *Human Factors*, **29(4)**, 383-394

Vicente, K.J. and J. Rasmussen (1992). Ecological Interface Design: Theoritical Foundations. *IEEE Transactions on Systems, Man, and Cybernetics*, **22(4)**, 589-606

NEW VISUALISATION TECHNIQUES FOR INDUSTRIAL PROCESS CONTROL

René van Paassen*

*University of Kassel, Human-Machine Systems, Mönchebergstraße 7, 34109 Kassel

Abstract. Using as an example a cement milling plant, this paper discusses the steps necessary for the development of a Man-Machine Interface on the basis of two novel techniques, Multilevel Flow Modelling (MFM) and Ecological Interface Design(EID). These techniques can be used to convey functional information in the interface, thus supporting the human operator in his goal-oriented behaviour. In a combination, the techniques are complementary rather than competing, where MFM provides support for failure analysis and EID provides support for fine-tuning and manipulation of the milling process.

Keywords. Process control, man/machine interfaces, computer interfaces, supervisory control, multilevel systems.

1. INTRODUCTION

The application of computer generated Man-Machine Interfaces (MMI) allows designers of these MMI's a great freedom over the layout and presentation of process data. Usually these displays follow the principle of one-sensor, one-indicator (Goodstein, 1981). Additionally they provide topological information, by showing the components and the connections between these components. They may also convey geographical information, i.e. show a resemblance to the layout of the plant and represent the components with icons that resemble the actual components in the plant (Larsson, 1992).

These conventional displays allow an easy recognition of the components in the plant and show the connections between these components. They do not provide clues about the purpose of these components. However, in his job the human operator (HO) must display a goal-oriented behaviour. With conventional displays the possible contributions of the components to the goals that must be achieved, i.e. the functional relations between the components of the interface, have to be learned and remembered by the operator.

Functional interfaces are types of MMI that support such goal-oriented behaviour. Examples of these interfaces are interfaces designed with *Multilevel Flow Modelling* (MFM) developed by Lind (1990) and *Ecological Interface Design* (EID) by Vicente and Rasmussen (1990).

These interfaces are based on means-ends views of the controlled plant. In these views the ends are the goals that must be accomplished, and the means are the functions that are available to reach those goals. When the means and ends are represented in a hierarchical structure, the means at the lowest level in this structure are functions implemented by the plant's components. These are the means to implement more abstract functions at higher levels in the means-ends hierarchy, and so on, up to the main goal of the plant. Multiple levels of means and ends may be necessary to represent a process.

This paper describes a part of the work for a Brite/EuRam fellowship, in which displays, for the example applications in the AMICA[1] project are developed and will be evaluated using MFM and EID.

In the following a short description of example application B of the AMICA project, a cement milling plant, will be given. Some designs for the MFM- and EID-based MMI for this plant will be discussed.

2. CEMENT MILLING PLANT

The raw material for cement milling is clinker. In the example plant two diggers can dig clinker from two different positions in the clinker store. The

[1] Brite/EuRam project 6126 AMICA, supported by the European Commission

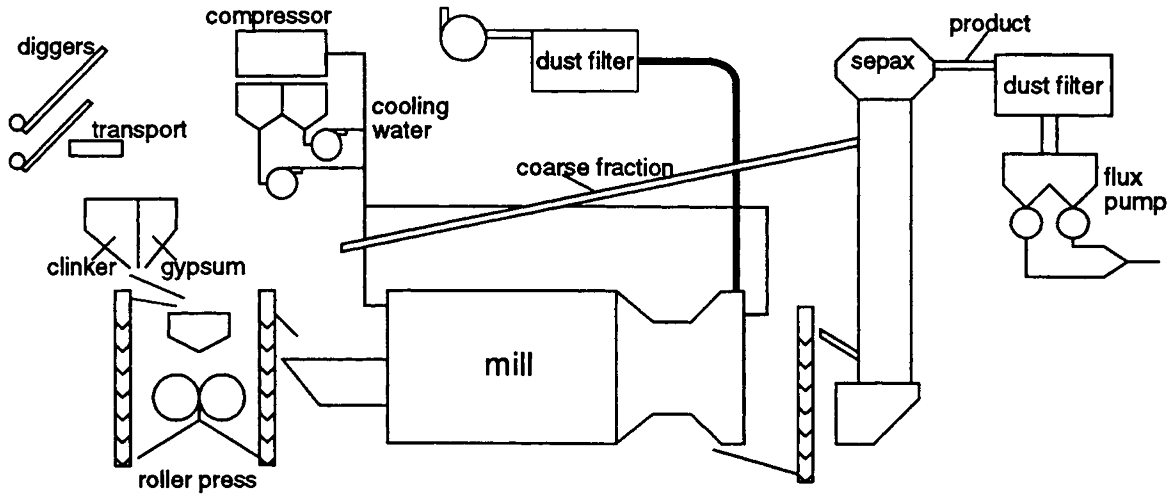


Fig. 1. A diagram of the cement milling plant.

clinker is transported with conveyor belts to the milling plant (Figure 1).

There gypsum is added to the clinker, and the materials are fed to a roller press, which crushes the clinker. Pressure and speed of the roller press have to be adjusted by the operator. Part of the output of the roller press is recycled through the press and a part is fed into the mill, where it is ground by the grinding media in the mill. The operator must adjust the feed rates of the different clinker types and the gypsum, pressure and speed of the roller press and its feedback ratio.

The output of the mill is fed into a SEPAX separator, which separates the coarse and fine fractions of the milled product. The coarse fraction of the product is fed back to the mill. The fine fraction of the mill output forms the desired product. It is lifted from the SEPAX by an airstream, separated in a dust filter and then transported to the product silos with a flux pump. The SEPAX also provides cooling of the milled product to approximately 70 °C, the maximum temperature at which the product may be stored.

In order to prevent spoiling of the product, the mill and the material in it have to be cooled to a temperature of approximately 115 °C. This is done by fine spraying of water into the mill. Water vapours and dust are taken out of the mill by a fan, and filtered in an electrostatic precipitator.

The display presently in use at the cement mill resembles Figure 1, and is an example of a topological display. It shows the components and their interconnections, but not their functions. In functional terms, the water injection system for cooling can be seen as a means to transport energy (heat) *from* the mill. The water injection system provides the support for this energy transport, i.e. it is a means for the "end" energy transport from

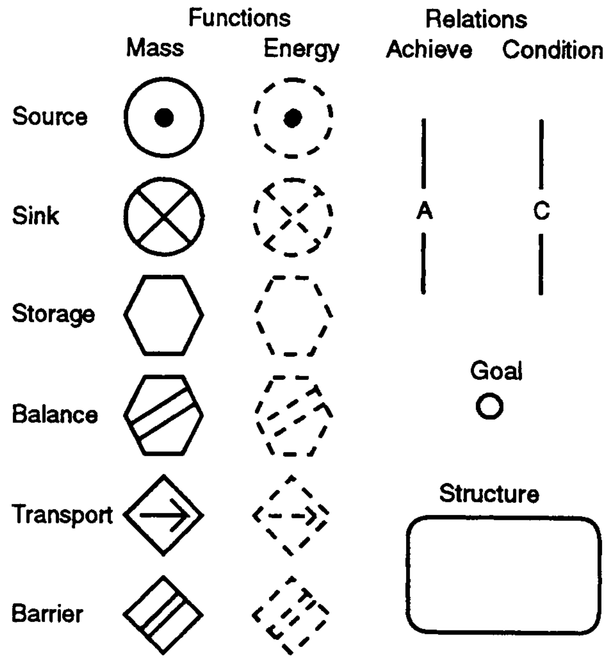


Fig. 2. MFM symbology, after (Larssen, 1993)

the mill. It is this type of information that can be presented with functional displays.

3. MULTILEVEL FLOW MODELLING

In MFM, a process is represented by its mass, energy and information flows. The symbology for this modelling is given in Figure 2. MFM is primarily a technique for modelling of a plant, but it is also possible to use an MFM model and the MFM symbology to build MMI interfaces.

The goals (ends) in the means-ends hierarchy used in multilevel flow modelling are achieved by the flow functions, which represent mass, energy or information flows in the plant. Flow functions are grouped into logical units, called structures. Usually a combination of flow functions in a struc-

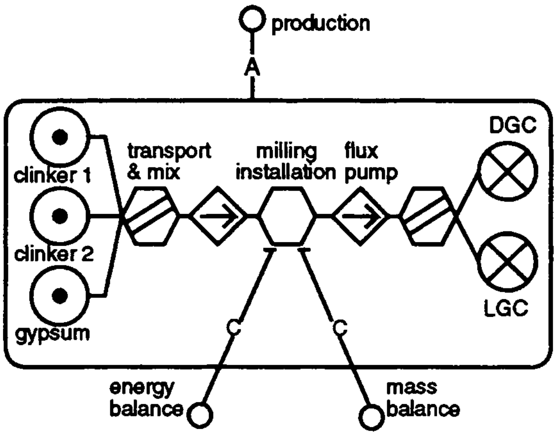


Fig. 3. MFM flow structure. The structure achieves the goal production, the operation of the mill depends on the sub-goals "mass-balance" and "energy balance"

ture forms the means to achieve one or more goals (ends).

The flow functions have to meet certain criteria, e.g. a level must be maintained in a storage function, or the flow of a transport function must be within certain bounds, in order to achieve a goal. At the lowest level in the means-ends hierarchy the flow functions usually correspond to single components in the plant, e.g. a transport function represents the transport of material through a valve or on a conveyor belt. At higher levels the flow functions may be more abstract. Usually the functions at these levels are established depending on goals lower in the means-ends hierarchy.

4. MFM MODEL FOR THE CEMENT MILL

For the construction of an MFM model, the goals in the system and the functions that are necessary to achieve these goals must be identified. The main goal in the example will be called "plant operation". As in most process control applications (Sassen *et al.*, 1994; Larsen, 1993; Lind, 1990) , three sub-goals can be identified:

- Production, of cement in this case.
- Safety and Environment.
- Economy.

The flow functions that achieve the cement production are given in Figure 3. This figure shows the three sources; the two clinker diggers and the gypsum supply. A balance function models the mixing of these components. The mill plant itself is modelled as a storage, because it can contain a variable amount of material being milled.

The output of the plant is transported to a balance, which corresponds to the valve by which the output silo can be selected. The first sink, is a silo, which contains the Lower Grade Cement (LGC), the second sink is a silo which contains the desired product, Desired Grade Cement (DGC).

Such a set of flow functions is called a flow structure. The conditions under which this structure achieves the production goal are:

- The input flow to the mill has an acceptable value.
- A proper mix between gypsum and clinker is achieved.
- The output flow has an acceptable value.
- The valve is set to silo with DGC.
- The quality of the clinker input is satisfactory.
- The quality (fineness) of the produced cement is satisfactory.

This structure is a very abstract description of the milling process. It is only fit for obtaining a global view of the plant state.

Those flow functions that can be directly controlled at this level will be marked in the operators display with a shaded edge. By clicking these functions, e.g. the balance that represents the silo valves, the operator can open a control pop-up to alter the setting of the valves.

The milling function itself, represented by a single MFM storage function in this example, is of course a complicated process. Its proper functioning depends on a two of conditions, i.e. sub-goals that must be met.

- Mass flow control. The level in the mill, input and output flows and recycle ratio have to be controlled. A proper mix has to be maintained for the input materials.
- Energy flow control. The proper amount of "milling energy" has to be added to the clinker, and this energy has to be taken out again by cooling of the mill and the finished product.

Below this level even more subgoals can be distinguished. For example for cooling the mill, a function in the energy flow control, water is sprayed in the mill, for which both water pressure and air pressure are needed. These structures and subgoals are not elaborated here.

5. ECOLOGICAL INTERFACE DESIGN

Ecological Interface Design is based on principles from Ecological Psychology (Gibson, 1979). The principal idea behind EID is that the affordances of the system, i.e. the ways in which the system can be manipulated to attain the system goals, are made visible in the display (Vicente and Rasmussen, 1990; Vicente and Rasmussen, 1992). As Rasmussen (1974) put it:

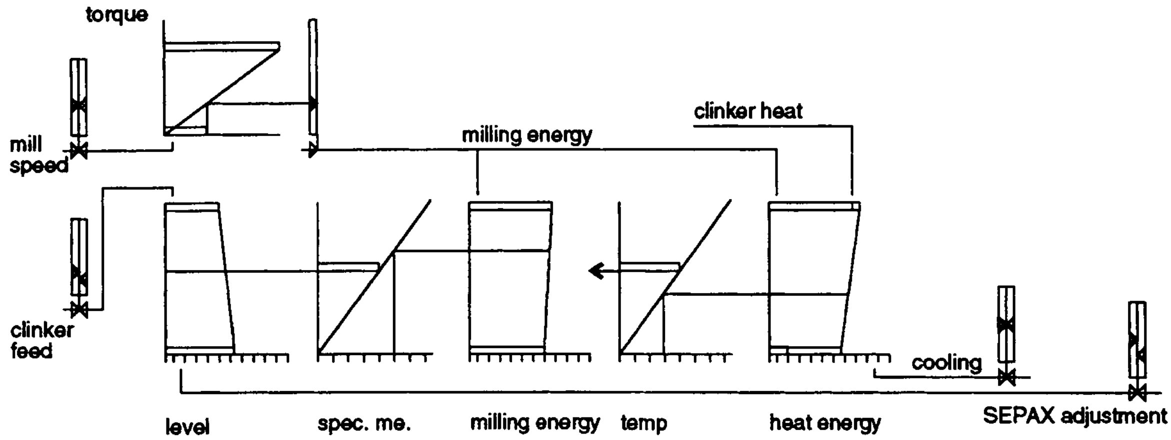


Fig. 4. Design for an EID interface for the energy and mass flow in the mill. The controllable variables (mill speed, clinker feed, cooling and SEPAX adjustment) are presented with control bars showing the setpoint and actual values.

> ...the display system therefore should be "transparent" and the physical process should be directly "touchable" on the control desk.

An example of such a display for a thermohydraulic process is given in (Vicente and Rasmussen, 1990).

The development of a display based on EID will be explained at the hand of an example, a display for the energy and mass management of the cement mill.

6. ENERGY AND MASS MANAGEMENT WITH EID

A disadvantage of MFM is that the energy and mass flow functions are represented in different structures. In many processes the operator has to maintain a balance between the mass and energy flows. The flow functions are quite useful when the process has to be checked for failures, but balancing mass and energy flows requires switching between two MFM structures.

In the cement milling process for example, the material level in the mill will be slowly increased until the level required for steady production is reached. During this process the energy flows, in the form of milling energy and water cooling, have to be repeatedly adjusted to the new level. Currently the operators deal with this by carefully changing only one variable at a time, and observing its effect before proceeding with other variables (Heuer *et al.*, 1993).

The design for the EID display for the energy and mass management of the milling installation, which in this case comprises the mill, its cooling system and the SEPAX separator, is given in 4. A characteristic of EID is that usually not simple measured values are indicated, with rote instruments, following the single sensor, single indicator principle, but also complex values may be presented. The leftmost graph on the bottom row in 4 is the mass balance graph. It provides combined information about the inflow of mass to the milling installation (horizontal bar at the top) the outflow (horizontal bar at the bottom) and the level of milled material. The slope of the graph indicates the trend of the level, using the principle of a funnel.

The input flow of material to the mill can be adjusted directly by the operator. The output flow can be influenced by the adjustment of the SEPAX rotor, but this will have consequences for the fineness of the ground material.

A second task that the operator faces is the application of energy by means of the motor of the mill and the removal of the heat from the mill and the milled material. The rightmost graph in 4 shows this energy input and output. Not all of the energy input and output can be individually controlled by the operator, part of it is determined by the temperature with which the material enters and leaves the mill.

The fourth graph in the bottom row shows the temperature of the material in the mill. It explicitly visualises the relation between the energy contents of the material and the amount of material in the mill, making it easier for the operator to judge the effect of his actions.

With the current MMI, operators closely watch the torque on the mill motor, which is an important status variable for the milling process. The torque, when multiplied by the rotation speed, gives the energy input to the mill. This is represented by the graph in the upper row of Figure 4. It can be said that for milling to a certain fineness a certain amount of energy has to be absorbed and released again by the ground material. The total amount of energy absorbed by the milled material is given in the middle graph of the bottom row. The output of this graphs is determined by the product outflow of the milling installation and the fineness of the product. The second graph in the bottom row shows the energy uptake per kg of material in the mill.

7. DESIGN OF THE MMI

A software library for evaluating the goals and flow functions of MFM models has been written in C++. Using this library, the MFM model for the cement mill will be implemented. On the basis of the sensor values of the cement mill, in this case generated by a simulator, the status of the flow functions and the goals in the means-ends tree is calculated. The alarms calculated for the MFM flow functions and goals will also be used for the EID displays.

The graphical presentation of the MFM and EID displays will be designed with the graphical tool GMS. The interface will permanently show a means-ends tree, which contains a breakdown of the goals, subgoals and functions in the plant. Three status conditions are indicated on these goal icons:

- A clear icon, showing the goal is met.
- An icon with a green exclamation sign, the goal is met, but probably will need attention soon.
- An icon with a red triangle, the goal is not met.

The rationale for the introduction of an "attention" stage is that experience with MFM displays has shown (Rauh, 1994; Sassen *et al.*, 1994) that, in order to give the operator the time to remedy an error, operators need an early warning. Using only one indication, and flagging this when the goal is not met, results in extra work, because in order to prevent this situation the operator will check the plant very frequently. The other solution, giving only an early indication when the goal is still met, leaves operators in doubt about when the goal is really not met.

Clicking a goal icon in the means-ends tree will open the associated MFM structure. When also an EID display is available for this goal, this can be reached by clicking an EID button in the window with the MFM structure. The "economy" goal and goals below it in the means-ends tree will directly open the associated EID display, because it is assumed that the operator will want to fine-tune the process.

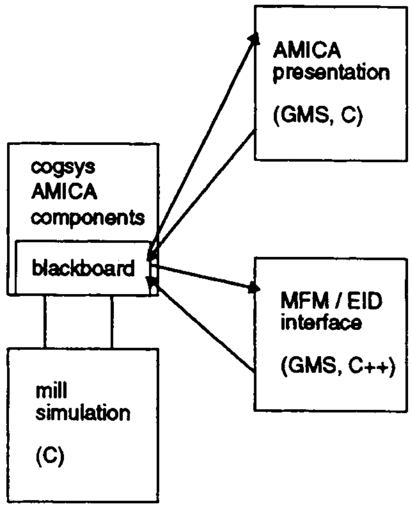


Fig. 5. Software components of the combined AMICA and MFM/EID system for the cement mill simulation. Communication is achieved by the COGSYS blackboard.

8. IMPLEMENTATION

The MFM and EID based display design will be evaluated with a simulator of the cement milling process, which is kindly provided by one of the industrial partners in the AMICA project, FLSA in Denmark. This simulator, written in ANSI C, will interfaces with other components in the AMICA system and with the MFM/EID software, via COGSYS, a tool for implementation of real-time knowledge based systems. The graphical display of both the MFM/EID interface and the AMICA interface is implemented with GMS.

An overview of the software components is given in Figure 5. The blackboard of COGSYS will be used as the communications medium. Output values of the simulation are written on the blackboard, these values can be used by both the AMICA interface and user support systems, written partly in COGSYS and partly in C, and by the MFM/EID interface. Likewise the control inputs for the simulation are written on the COGSYS blackboard, and read by the simulator.

9. DISCUSSION

The combination of an MFM display and an EID based display appears to be a feasible one. The development of EID displays profits from the efforts of the development of the MFM model, since the controls and graphs in EID correspond to transport functions, balances, and storage functions in MFM.

The two techniques are complementary rather than competitive. For the diagnosis of malfunctions and alarms in the plant MFM offers a consistent and efficient solution. The fine tuning of process and the interaction with process dynamics can be more efficiently done with EID, since operations on the plant are easily visible.

A combination with topological displays is preferable, since this will increase operator acceptance. In this view the combination in the AMICA project, where an MMI with advanced display facilities and operator support is developed, is optimal. [2]

10. REFERENCES

Gibson, J.J. (1979). *The Ecological Approach to Visual Perception.* Houghton-Mifflin. Boston.

Goodstein, L.P. (1981). Discriminative display support for process supervisors. In: *Human Detection and Diagnosis of System Failures* (J. Rasmussen and W.B. Rouse, Eds.). pp. 21–36. Plenum Press. New York.

Heuer, J., S. Borndorff-Eccarius and Averbukh E.A. (1993). Task analysis for application b. AMICA BRITE/EURAM Project 6126 IR1-04. University of Kassel, Laboratory for Human-Machine Systems.

Larsen, M.N. (1993). Modelling start-up tasks using functional models. Technical Report EUR 15027 EN. Commission of the European Communities, JRC–Ispra.

Larsson, J.E. (1992). Knowledge-Based Methods for Control Systems. PhD thesis. Department of Automatic Control, Lund Institute of Technology.

Lind, M. (1990). Representing goals and functions of complex systems - an introduction to Multilevel Flow Modelling. Technical Report 90-D-381. Institute of Automatic Control Systems, Technical University of Denmark.

Rasmussen, J. (1974). The human data processor as a system component: Bits and pieces of a model. Technical Report Risø-M-1722. Danisch Atomic Energy Commission. Roskilde, Denmark.

Rauh, J. (1994). Erstellung einer Benutzeroberfläche mittels Multilevel Flow Modelling. in German: Construction of a user interface using Multilevel Flow Modelling. Master's thesis. University of Kassel, Laboratory for Human-Machine Systems.

Sassen, J.M.A., E.F.T. Buïel and J.H. Hoegee (1994). A laboratory evaluation of a human operator support system. *Int. J. Human-Computer Studies* **40**, 895–931.

Vicente, K.J. and J. Rasmussen (1990). The ecology of human-machine systems ii: Mediating "direct perception" in complex work domains. *Ecological Psychology* **2**(10), 207–249.

Vicente, K.J. and J. Rasmussen (1992). Ecological interface design: Theoretical foundation. *IEEE Transactions on Systems, Man and Cybernetics* **22**(4), 589–606.

[2] The presented work has been carried out as part of the European Community research training project "Visualisation of Process Means-Ends Hierarchies for Advanced Man-Machine Interfaces" (ERBBRE2CT943085). The work has been done in cooperation with the partners in the BRITE/EURAM II Programme in Project 6126 (AMICA: Advanced Man-Machine Interfaces for Process Control Applications). The consortium partners are: CISE (Italy), ENEL (Italy), FLS Automation A/S (Denmark), University of Kassel (Germany), Marconi Simulation and Training (Scotland).

A PROPOSAL TO DEFINE AND TO TREAT ALARMS IN A SUPERVISION ROOM

B. RIERA, B. VILAIN, L. DEMEULENAERE and P. MILLOT

* *University of Valenciennes, LAMIH, Le mont Houy, BP 311, 59304 Valenciennes, FRANCE*

Abstract: Automation offers the possibility of building production systems more and more complex. This evolution leads to a modification of human operator's activities which sets out problems, particularly when a failure occurs in the process. The first part of the paper presents, after the human operator's work and man-machine interfaces in a supervision room, a synthesis of different problems linked to alarms. The second part of the paper is a proposal to design and to treat alarms. The goal is to design and display "educational" alarms, well defined and well treated, giving to the human operator the means to perform correctly his work. On that subject, the Multilevel Flow modelling (MFM) for the functional analysis and methods from the dependability science like FMEA (Failure Mode and Effects Analysis) are interesting.

Keywords: Man-machine systems, ergonomics, alarms, reliability analysis.

1. INTRODUCTION

Automation offers the possibility of building production systems more and more complex. This evolution leads to a modification of human operator's activities which sets out problems, particularly when a failure occurs in the process.

The first part of the paper presents, after the human operator's work and the man-machine interfaces in a supervision room, a synthesis of different problems linked to alarms. The second part of the paper is a proposal to design and to treat alarms. The goal is to design and display "educational" alarms, well defined and well treated, giving to the human operator (HO) the means to perform correctly his work. On that subject, the Multilevel Flow modelling (MFM) developed by Lind (1990) for the functional analysis and methods from the dependability science like FMEA (Failure Mode and Effects Analysis) are interesting.

2. THE HUMAN OPERATOR IN A SUPERVISION ROOM

The supervision of industrial processes includes a set of tasks aiming at controlling a process and at supervising its running (Millot, 1988). The control consists in acting on the process ; so, it is a top-down flow of information which acts on the lower levels (Benzian, 1994). On the contrary, the supervision is a bottom-up flow of information of which sources are the signals sent by the process. Control and supervision tasks require high level of knowledge from the human operator and can be grouped in three classes (adapted from Rouse, 1983) :

- The transition tasks corresponding to a change of the running mode of the process (for instance, the start-up or the stop of the production system). In this case, the operator, most of the time, has to perform procedures already defined.
- The control tasks and the supervision tasks of the normal running of the process. The operator supervises the process and he tries to optimize the running by the means of adjusted tunings. Lejon (1991) points out that these tasks are going to disappear because of optimization algorithms.
- The detection of failures tasks, the diagnosis tasks and the resumption tasks. The HO has in a first time to detect the presence of an anomaly by the means of alarms displayed or the trends of some variables. In a second time, the operator has to diagnose the state of the process. This consists in finding the initial cause of the failure but also the effects on the process. It is important to note that the level of abstraction and knowledge of the supervision operator is different from a maintenance operator's one. Indeed, if a process is considered as a structured set of components enabling to realize functions, then the maintenance operator's work consists in fixing the physical structure in order to have its normal running back. The notion of physical structure is, in this case, microscopic and means the machines. On the other hand, the supervision operator has to determine the physical structure which does not run well. In this

case, the notion of physical structure is macroscopic. After having diagnosed the state of the process, the HO has to compensate or to correct the defect. These tasks depend a lot on the nature of the failure. If components are broken, the HO, if he can, would start up other equipments in order to isolate the failing sub-system. If he can not, he would continue to control the process in a debased mode or lastly he could decide to stop the process in order to fix devices.

So, automation and centralization of information in a supervision room have considerably modified human operator's work. One can note that when everything is going well, that is to say when there is a normal running of the process, the HO in front of his screen, far from the process, can damage his mental representation of the structure and the functions of the process. The associated risk is a drop of operator's performance when an abnormal running occurs.

As it has been seen previously, the essential work of an operator in a supervision room is to get at each instant a correct view of the production system. By the past, when there were not supervision room and complex control loops, the operator was in front of his machine. Hence, he really lived the running of the machine and he saw continuously the physical structure. HO controlled with his eyes and not with figures. Being close to the machine, he could detect quickly failures. More, the operator fetched information with his eyes (a level of a tank for instance), it was no use to cast doubt over their validity. Nowadays, in a supervision room, information about the process come from sensors which can be out of order. It is easy to understand the difficulty for an operator in a supervision room to diagnose the state of the process with information being able to be erroneous. For that, the operator has to confirm data and he must have an analytic mind.

Studies realized in supervision room for industrial processes, show that most of dysfunctions come from :

- false information coming from sensors,
- locking of automatic floodgates,
- wrong understanding of the system by the supervisor operator, i.e. wrong control set-points, wrong decisional thought process or wrong interpretation of control loops,
- equipments out of order.

It can be noticed, that the supervision room is a crucial informative node. The maintenance staff, the automation staff or the management staff are interested in information centralized in the supervision room. The HO must, in fact, optimize 3 criteria when he controls the process :

- The production criterion. The first objective of the production means consists in the products transformation and the added values creation. This objective must be managed in guaranteeing the reliability and the availability of the equipments.
- The security criterion. This criterion is the most important because the supervision HO must guarantee, first of all, the security of the human beings and of the equipments.
- The economical criterion. The production must be the cheapest as possible.

One can understand the importance of the man-machine interface in the supervision room. The man-machine interface must permit the HO to have a right vision of the process state. From the information, which are presented on the screen, the HO must conjure up a correct structural and functional model of the installation. However, it will be false to believe, that the quality of the interface can be a stopgap. HO must be high motivated and efficient. A well defined work organization, that can allow the operator to maintain and improve his knowledge and a robust and stable process must be exist too. It is all of these parameters will guarantee a good performance of the global Man-Machine System.

3. THE DESIGN OF SUPERVISION IMAGERIES

The main objective of the supervision imagery is to give means to the HO to control and supervise the process. The principal difficulties come from the obligation to synthesize a large number of information on the screen. The common design mistake consists in the presentation of the maximum of information, in order to be sure that nothing has not been forgotten (Notte, 1986).

Vittet (1981) and Kolski (1993) proposed recommendations, which can be applied to the design of supervision imagery :

- To supply a global view of the installation and its running to the HO. He must access to these pertinent information, without too much reasoning.
- To supply information concerning the evolution of the installation state to the HO. In particular, the changes of state must be presented (by means of alarms for example).
- To supply to the HO information which permit to control quickly results of his action.
- To allow the HO to drive away the different levels of process abstraction.

Nowadays, one can notice a tendency to propose support systems to the HO, in order to assist him. These can help the HO in his decisional thought process (detection help, problem solving help and action help). Usually, the decision support systems are based on artificial intelligence techniques. It looks paradoxical to graft additional systems to the supervision system in order to help the HO. In fact, it will be more judicious to supply the HO with data and information, which he really needs to achieve his work. We think that support systems do not have to be seen like system replacing the decisional thought process of the HO, but like toolbox, which facilitates the HO work.

Very often, supervision imageries are designed by the automation people who are not supervision operators. The synoptic views often show the structure of the installation. The presented information are not always appropriated with the real operator needs. It is why several experiments to have the design supervision imagery made by the operators (Kolski, 1992) have been carried out. This solution is attractive, but it is not obvious to have the supervisors expressed their needs. The imagery designers are often guided by their good sense, their experience and the used material. Trends and regulators are often displayed on control screen. The changes of process state are displayed by alarms which appear on control screen. The next paragraphs deal especially with problems of alarms.

4. ALARMS

In theory, an alarm is a signal, which informs the operator of a danger. In a practical view, in supervision room, alarms amalgamate a set of very different elements. The classification of alarms suggested by Wanner (1987) is composed of two categories :

- The not-programmed alarms, i.e. alarms, which are not defined at the time of the supervision system design, but these are used by the HO, that goes round the site: abnormal noise, smoke, steam, explosion, ... In this article, theses aspects will not be developed.
- The programmed alarms, i.e. alarms, which are defined at the time of the supervision system design.

This second category can be divided into 2 groups of alarms :

- The breakdown alarms, which indicate a failure. Without these alarms, the detection of the failure will be accessible with the observation of several variables and an important cognitive reasoning.
- The process alarms, which show an abnormal performance of the process.

In fact, the first alarms correspond to failures of the physical components of the process (i.e. the structure). These alarms are defined by the automation team. For instance, motor disjunction, floodgate locking can be quoted. In a lot of supervision rooms, the alarms are not sufficiently commented. The term "default" usually appears. This involves interpretation problems for the HO. The first characteristic of an alarm system is to provide commented alarms. Hence, alarms, like "default of #i floodgate" should not exist and must be replaced by alarms "no opening" or "no closing" of the #i floodgate. As well, the unique piece of information "default of #i motor stopping" can be a source of confusion for the HO. It must be indicated if it is a “normal” stop activated by a plant operator or if it is really a motor failure.

The second alarms are more difficult to define. In fact, they correspond to an abnormal behaviour of the process. These alarms represent the speech of an experimented operator (e.g. the value is too high compared to the set point, this rate of flow is important). A priori, these alarms are complex but they permit the operator to anticipate possible dysfunctions of the process. These alarms are mainly functional because it is the observation of a function which will permit to detect a possible process dysfunction. Generally, the threshold overtaking of one or several variables triggered off the alarms. These thresholds are difficult to define. On the one hand, if the threshold is too low, the alarm will be always activated and will be sensitive to noise. On the other hand, if the threshold is too high, the alarm will trigger off late and the operator will not be able to anticipate. An another solution consists in creating alarms with a combination of conditions. But, the design of theses alarms is more difficult and requires a deeper knowledge of the process, which is generally only known by the process engineer, the expert of the process. Even if the exhaustive list is very difficult (almost impossible) to define, this approach seems attractive.

So, it is essential, before processing the alarms, to well define them. This preliminary stage is absolutely necessary.

All of these evoked points show that the alarm processing system must be educational. The alarms must be easily understandable and usable by the operators. In addition, alarms enable to understand better the process. Indeed, The HO experiences partly come from incidents. So, the alarm processing system must facilitate the HO training, it must allow the HO to exploit as much as possible the dysfunction appearance. That permits the conceptualization of the process, its functions as well as its structure. So, the alarms must be a source of explanations.

Today, alarms displaying and processing system do not respect these objectives for several reasons which are presented now.

5. THE INCONVENIENCES OF CLASSICAL ALARMS PROCESSING SYSTEMS

Numerous defaults of present alarms processing systems exist :

- Lack of taking into account the process : all the alarms are not significant of a breakdown, or more generally of an abnormal state. It depends on the situation of the process. In particular, during the specific running modes of the process (start-up, shut-down, maintenance, cleaning, ...), the evolution of certain parameters has not the same signification. The different running modes must be taken into account in order to have always alarms associated to an abnormal situation. Nowadays alarms are generally designed in taking into account a local context. So, without taking into account all the contexts, some alarms will go off, whereas they do not, and will create trouble for the operator, who will have to find significant alarms. In fact, there is a lack of alarms reliability (De Keyser, 1980).

- Lack of taking into account the events causality : all the process variables can be linked together by a relation of cause to effect. Indeed, the diagnostic is a causal process. A fault cause involves fault mode which involves effects on the process. With regard to different dynamics of the process and the alarms thresholds placements, some alarms appear without respecting the causal link (the effect appears before the cause). If causal link are not respected, the HO can get a wrong mental image of the process working and he can have difficulties to diagnose a default. Today alarm processing systems do not give an ordering alarms list. So, the functional localisation of the first default is not easy. If a direct link with the synoptic view does not exist, the structural localisation is not easy too.

- Specific problems of alarms activated by threshold overrange : flopping alarms or fugitive alarms are well known phenomena, which can be solved by means of temporal filters.

- Alarm processing systems are not very suitable with HO characteristics. He is often reactive on alarms and is waiting alarms, instead of anticipating the situations. At last, due to limited human capabilities of information processing, the operator can not faces up easily to alarms cascade.

Most of the time, all these problems, due to wrong alarm design, can have negative consequences. In front of repetitive alarms, the operator can acknowledge the alarms without analysing them, or even can suppress them. That could be very dangerous for the security and the production.

6. SPECIFICATIONS FOR THE DESIGN AND THE PROCESSING OF ALARMS

The set of evoked points concerning the supervision operator and alarms brings about these following specifications for alarms design, alarms processing and alarms display.

This tool must be educational. So, it is necessary :

- To well define alarms. The alarms can be divided into two categories : breakdown alarms and process alarms. It is necessary to comment them for the HO.

- To show the causal relation between the different variables. This representation permits, on the one hand, to have a correct functional vision of the process and on the other hand, to get a direct vision of source alarms. Methods using causal graphs (Montmain, 1994) can help the HO in his diagnosis task, by displaying source alarms. Even if these methods permit to identify the first effects on the process, and not the breakdown itself, they help really the HO.

- To classify alarms according to the gravity and to show the consequences according to 3 criteria : "production, security, economical".

- To define alarms, which are specific to the different running modes of the process.

- To enable the operator to work at different levels of abstraction with alarms.

The foreseen reliability methods present a great interest for the design of breakdown and process alarms.

7. ALARM DESIGN AND FMEA

Some methods from dependability science can be used to help the design of alarms. The methods, like FMEA (failure modes and effects analysis) (Villemeur, 1988), consist in a qualitative and quantitative evaluation of potential failures of a system and in the study of effects (or consequences) for the user. The system can be a product (product FMEA) or a process (process FMEA). The FMEA helps to improve reliability and maintainability of a device, putting out the risk points in order to reduce them by appropriated measures. In theory, the FMEA is made as soon as possible during the design stage in order to have time to make corrective actions which will prevent the breakdowns. The FMEA is managed by a team composed of persons, having different competences and knowledges about the system.

We propose to use the principle of breakdown evaluation of the method in terms of modes, causes and effects in order to define a specific FMEA for the supervision operator. This FMEA, which is called INHO FMEA (that is to say information need of human operator FMEA) must be a real help for the alarms design. The INHO FMEA is presented now.

The functional analysis is the preliminary stage of all FMEA. This analysis permits to get a process decomposition in several sub-sets performing specific functions (see Fig. 1).

A possible approach for the functional analysis could be MFM (Multilevel Flow Modelling) developed by LIND (1990).

The aim of MFM is to supply a systematic basis to build models of complex systems. The decomposition "means - ends" (means to obtain the ends) and "whole - parts" is used. The system is described in terms of goals and functions. The goals characterize the objectives to be attained by the system. The goals are achieved by the functions and functions are supported by the physical components of the system. By analogy with the 3 criteria, that HO must optimize, 3 kinds of goal exist (LIND, 1990) : production goal, security goal and economical goal.

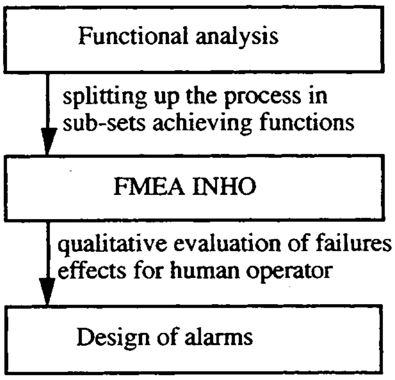


Fig. 1. From the functional analysis to the design of alarms.

The goal formalization can be qualitative or quantitative.

The functions represent the means which are used to achieve the goal. LIND (1990) defines 6 basic functions : storage, balance, transport, barrier, source and sink. All these functions permit, a priori, the description of the mass or energy structures achieving the goals. Each basic function can be described with a set of components. This kind of functional analysis is very suitable for the INHO FMEA (see Fig. 2).

Goals	Production	Security	Economic
Sub-goals			
Functions			
Set of components			
Components			

Levels of abstraction

Fig. 2. Table 1 for functional analysis.

The next step consists in filling in the table presented in the figure 3. This table contents, on the one hand, all the system components and on the other hand, basic functions, which define mass and energy structures. This table will be the support of the INHO FMEA.

The objective of the INHO FMEA is relatively different from a classical FMEA. Indeed, estimated breakdown analysis is not performed in order to guarantee the availability, the reliability and maintainability of the production system by the means of corrective actions. The aim is to have a Human-Operator centred approach, in order to define his information needs, including the design of breakdown and process alarms of the system.

GOALS	G0				G1					G2				
Functions / Components	f01	f02	...	f0j	f11	...	...	...	f1k	f21	...	...	...	f2l
C1	X			X					X					X
C2		X		X	X									
C3	X				X				X	X				X
C4	X								X					X
...														
Ci		X		X	X					X				
Ci+1	X								X					X
...														
Cn		X		X										X

Levels of abstraction

Fig. 3. Table 2 for functional analysis.

The components include physical structure elements and transported material. For example, in the case of a central heating system, the components are : the oil tank, the pump, the boiler, the heater, and so on, but also water and fuel oil. One can notice that the table 2 can be directly filled in from the MFM model.

The next step consists in the filling up of the INHO FMEA table (see Table 1). This table is composed of 9 columns, which are presented now :

- Functions : all the functions, defined by the functional analysis, must be studied.
- Components : for each function, the possible failures of components and materials which support the function, must be studied.
- Failure modes : they correspond to the way the components stop or have an abnormal functioning.
- Possible causes of the breakdown : for each component, the initial anomaly which could lead to the breakdown is notified. It is important to precise if it is a material cause or an internal failure of the component, in order to distinguish breakdown and process alarms.
- Effects : no achieved goals have to be specified.
- Detection : in this column, one indicates what are a priori existing means or means having to be implemented in order to enable the operator to detect the failure.

- Severity : the breakdown gravity is estimated in terms of production, security and economical consequences. For instance, a range from 1 to 10 can be used. A threshold, representing the obligatory detection for the HO, can be decided.

- Breakdown alarms : one indicates the possibility or not to have the breakdown information (if there is an internal component breakdown). If these information exist, the alarm label is written in order to be usable and comprehensive for the HO. In the other case, one indicates the directly influenced alarm (generally process alarms) permitting to infer the breakdown alarm.

- Process alarms : one indicates the possibility or not to have the process information. If the information are available the numerical formula will trigger the alarm must be defined.

This INHO FMEA table, associated with functional analysis based on MFM, contains a lot of information. This table can also be used to study the alarms cascade and the influences between alarms. At the moment, this method of designing alarms is being tested on practical examples. However, this method must be fined down and must be improved in order to take into account the different running modes of the process.

8. CONCLUSION

In this paper, the importance of the correct alarm design has been noticed. The alarms are classified into 2 groups : breakdown alarms, which correspond to internal failures of components, and process alarms. Some specifications to design a really performing and educational alarm processing and display system have been proposed. An alarm design method using inductive method from the dependability science (the FMEA) has been proposed. The FMEA is suitable for the definition of the operator information needs and the design of alarms. The preliminary functional analysis based on MFM seems attractive.

9. REFERENCES

Benzian, Z., J.L. Ermine, C.M. Falinower and B. Bergeon (1994). Modélisation des connaissances et spécification de systèmes de supervision de centrales thermiques. Journée d'Etude S3 (Sûreté, Surveillance, Supervision) du 8 décembre. Stratégies de conduite en présence de défaillances.

De Keyser, V. (1980). Etude sur la contribution que pourrait apporter l'ergonomie à la conception des systèmes de contrôle et d'alerte dans les industries de transformation. CCE, rapport S/79/45, Luxembourg.

Kolski, C. (1993). *Ingénierie des interfaces homme-machine : conception et évaluation.* Hermès, Paris.

Kolski, C. and D. Warroux (1992). Une expérience de conception, de réalisation et d'évaluation par des conducteurs de leur propre synoptique de conduite de réseau gazier. Le travail humain, tome 55, n°4.

Lejon, J.C. (1991) *L'évolution de la conduite sur S.N.C.C.* Edition Dunod, Collection point de repère.

Lind, M. (1990). Representing goals and functions of complex systems. An introduction to Multilevel Flow Modelling. ISBN 87-87950-52-9.

Notte, D. (1986). Ergonomie, applications informatiques dans les processus industriels continus : le point de vue des opérateurs. L'homme et l'écran : aspects de l'ergonomie en informatique, Actes du congrès de Nivelles : l'ergononomie en informatique. Editions de l'Université de Bruxelles, Belgique.

Millot, P. (1988). *Supervision des procédés automatisés et ergonomie.* Hermès, Paris.

Montmain, J. (1994). Raisonnement approximatif et graphes causaux pour la détection et la localisation de défaillances. Journées d'Etude S3 (Sûreté, Surveillance, Supervision) des 17 et 18 novembre. Détection et localisation de défaillances.

Rouse, W.B. (1983). Models of human problem solving : detection, diagnosis and compensation for systems failures. Automatica, vol 19, n° 6.

Villemeur, A., (1988) Sûreté de fonctionnement des systèmes industriels *: fiabilité, facteur humain, informatisation.* Ed Eyrolles, Paris.

Vittet, J. (1981). Eléments pour une conception intégrée des salles de contrôle. Thèse de docteur ingénieur, Université Scientifique et Médicale de Grenoble.

Wanner, J.C. (1987). Facteurs humains et sécurité. Séminaire : Erreurs humaines et automatisation. Ecole Nationale de l''Aviation Civile, 19-20 Mai 87.

Table 1 FMEA INHO

FUNCTIONS	COMPONENTS	FAILURE MODES	POSSIBLE CAUSES	EFFECTS	DETECTION	SEVERITY P	S	E	BREAKDOWN ALARMS	PROCESS ALARMS
F01	C1									
	...									
...										

REINFORCEMENT LEARNING AND DYNAMIC PROGRAMMING

Andrew G. Barto*

**University of Massachusetts, Department of Computer Science, Amherst, MA 01003, USA*

Abstract. Reinforcement learning refers to a class of learning tasks and algorithms based on experimental psychology's principle of reinforcement. Recent research uses the framework of stochastic optimal control to model problems in which a learning agent has to incrementally approximate an optimal control rule, or policy, often starting with incomplete information about the dynamics of its environment. Although these problems have been studied intensively for many years, the methods being developed by reinforcement learning researchers are adding some novel elements to classical dynamic programming solution methods. This article provides a brief account of these methods, explains what is novel about them, and suggests what their advantages might be over classical applications of dynamic programming to large-scale stochastic optimal control problems.

Key Words. Adaptive control, dynamic programming, function approximation, large scale systems, learning control, Monte Carlo method, optimal control, stochastic approximation, stochastic control.

1. INTRODUCTION

The term reinforcement comes from studies of animal learning in experimental psychology, where it refers to the occurrence of an event, in the proper relation to a response, that tends to increase the probability that the response will occur again in the same situation (Kimble, 1961). Although the specific term "reinforcement learning" is not used by psychologists, it has been widely adopted by theorists in engineering and artificial intelligence to refer to a class of learning tasks and algorithms based on this principle of reinforcement (e.g., (Minsky, 1961; Mendel and McLaren, 1970)). A reinforcement learning problem is usually formulated mathematically as an optimization problem with the objective of finding an action, or a policy for producing actions, that is optimal according to a given objective function.

Receiving the most attention in recent research are reinforcement learning problems in which the objective function measures the learning agent's behavior as it extends over time. In these problems, the relevant consequences of an action may not be available immediately after that action is taken. Recent research models these problems as stochastic optimal control problems in which the controller (the learning agent) has to find (or, more realistically, has to approximate) an optimal feedback control rule (a policy for acting), usually starting with incomplete information about the dynamics of the controlled system (the agent's environment). Although these problems have been studied intensively for many years, the methods being developed by reinforcement learning researchers are adding some novel elements to classical dynamic programming (DP) solution methods. Specifically, reinforcement learning algorithms based on DP all involve *interleaving* steps related to the operations of various DP algorithms with steps of real or simulated control. In other words, a DP algorithm is not used stictly as an off-line algorithm for designing a policy before that policy is used for control. Instead, DP acts as the organizing principle for incrementally compiling information acquired during real or simulated control experience.

This article provides a brief account of these methods, explains what is novel about them, and suggests what their advantages might be over classical applications of DP to large-scale stochastic optimal control problems. This account, which has been assembled by many researchers over roughly the last five years, represents our current state of understanding rather than the intuition underlying the origination of these methods. Indeed, DP-based learning originated at least as far back as Samuel's famous checkers player of the 1950s (Samuel, 1967; Samuel, 1959), which made no explicit reference to the DP literature existing at that time. Perhaps the earliest connection between learning and DP was suggested in 1977 by Werbos (1977), and much of the current interest is attributable to Watkins' 1989 dissertation (Watkins, 1989). Additional information on this class of algorithms can be found in (Barto, 1992; Barto *et al.*, 1994; Sutton, 1992; Werbos, 1992).

We begin with a brief description of the conventional DP approaches to these problems, restricting attention to finite Markov decision processes (MDPs), a formalism on which most of the rigorous theory of reinforcement learning relies. The principles of reinforcement learning, however, are by no means restricted to this class of problems.

2. THE CONVENTIONAL APPROACH

2.1. *Finite Markov Decision Processes*

At each discrete time step a controller observes a system's state x, contained in a finite set X, and selects a control action a from a finite set A_x. A reward with expected value $R(x,a)$ is delivered, and the state at the next time step is y with probability $p^a(x,y)$. A (stationary) policy $\pi : X \rightarrow \cup_{x \in X} A_x$ specifies that the controller executes action $\pi(x) \in A_x$ whenever it observes state x. For any policy π and $x \in X$, let $V^\pi(x)$ denote the *expected infinite-horizon discounted return* from x given that the controller uses policy π. Letting r_t denote the reward at time t, this is defined as:

$$V^\pi(x) = E_\pi \left[\sum_{t=0}^{\infty} \gamma^t r_t | x_0 = x\right], \quad (1)$$

where x_0 is the system's initial state, γ, $0 \leq \gamma < 1$, is a factor used to discount future rewards, and E_π is the expectation assuming the controller always uses policy π. It is usual to call $V^\pi(x)$ the *value* of x under π. The function V^π is the *value function* corresponding to π.

This is a finite, infinite–horizon, discounted MDP. The objective is to find an optimal policy, i.e., a policy, π^*, that maximizes the value of each state x defined by (1). The unique *optimal value function*, V^*, is the value function corresponding to any optimal policy. Additional details on this and other types of MDPs can be found in many references (e.g., (Bertsekas, 1987; Ross, 1983)).

2.2. *Dynamic Programming*

In the absence of structure beyond that described above, DP provides the only exact solution methods for MDPs short of an exhaustive search of policy space. *Value iteration* is the DP algorithm that successively approximates V^* as follows. At each iteration k, it updates an approximation V_k of V^* by applying the following operation for each state x:

$$V_{k+1}(x) = \max_{a \in A_x} [R(x,a) + \gamma \sum_{y \in X} p^a(x,y) V_k(y)]. \quad (2)$$

We call this operation a "backup" because it updates a state's value by transferring to it information about the approximate values of its possible successor states. Applying this backup operation once to each state is often called a *sweep*. Starting with an arbitrary initial function V_0, the sequence $\{V_k\}$ produced by repeated sweeps converges to V^*. When V^* is approximated to a desired accuracy, an optimal policy is taken to be any policy that for each x selects an action that realizes the maximum on the right–hand side of (2).

In the asynchronous version of value iteration (Bertsekas, 1982; Bertsekas, 1983), a backup can be applied at any time to any state using whatever values happen to be available for the successor states. As long as the value of each state continues to be backed up, this also converges to V^*. Thus, systematic sweeps of the state set are not necessary.

Policy iteration produces a sequence of policies that converges to an optimal policy. It starts with a an arbitrary initial policy, π_0, and computes its value function, V^{π_0}, then defines a new policy, π_1, that is optimal with respect to V^{π_0}. The process repeats starting with π_1. Each policy generated is either an improvement over the preceding policy, or is an optimal policy, π^*. Policy iteration's computational bottleneck is the necessity to evaluate a policy at each step, which requires solving a system of $|X|$ linear equations.

Although DP algorithms are polynomial in the number of states, they can be impractical because many problems give rise to very large state sets. For example, if X is a discretization of a multi–dimensional continuous space, its size is an exponential function of the dimension (Bellman's "curse of dimensionality"). Performing even a single value iteration sweep, or a single policy evaluation step of policy iteration, is often impractical. In fact, to exactly compute a backup operation (2) for a single state can itself be impractical due to the required sum over X and/or the max over A_x. It is also clear that these algorithms require a complete model of the MDP in the form of knowledge of the transition probabilities, $p^a(x,y)$, and the reward expectations, $R(x,a)$, $x, y \in X$, $a \in A$.

2.3. *The Case of Incomplete Information*

In the absence of a complete and accurate model of the MDP (the case of incomplete information), perhaps the most common approach (see, e.g., (Kumar and Varaiya, 1986; Kumar, 1985)) is to try to identify the unknown system and execute a DP algorithm under the assumption that the system model is correct (the certainty equivalence approach). This leads to the class of indirect adaptive control methods which at each time step of on-line interaction with the unknown system: 1) udpate the system model using current observations, 2) run a DP algorithm using the latest model, and 3) select a control action on the basis of the DP algorithm's results, taking into account the necessity to maintain behavioral variety (to "explore") due to the conflict between identification and control.

This approach thus builds conservatively upon the results for the complete information case, producing methods whose main attraction is theoretical tractability: they are practical only for very small problems. It is often not feasible to execute a DP algorithm *once*, let alone at each time step of control. Mitigating this somewhat is the fact that the number of iterations of, say, policy iteration, tends to be small if the initial policy is the final policy produced at the preceding time step. However, even in this case, at least one policy evaluation step is required, which can be too time-consuming to accomplish before the next control decision must be made, and can often be too time-consuming to accomplish under any circumstances.

3. APPROXIMATING DYNAMIC PROGRAMMING

The computational difficulties of solving large MDPs, with complete or incomplete information, are well-known, and many methods have been proposed for approximating their solutions with less effort (e.g., (Jacobson and Mayne, 1970; Norman, 1972)). To the best of this author's knowledge, DP-based reinforcement learning methods are novel in their combination of Monte Carlo, stochastic approximation, and function approximation techniques. Specifically, these algorithms combine some, or all, of the following features:

1. Avoid exhaustive sweeps by restricting computation to states on, or in the neighborhood of, multiple sample trajectories, either real or simulated (the Monte Carlo aspect).
2. Simplify the basic DP backup by sampling: for example, estimate the expected values on the right-hand side of (2) by sampling from the appropriate distributions (the stochastic approximation aspect).
3. Represent value functions and/or policies more compactly than lookup-table representations by using function approximation methods, such as neural networks.

In what follows, we elaborate each of these properties by describing several reinforcement learning algorithms.

3.1. *Avoiding Exhaustive Sweeps*

The reinforcement learning algorithm closest to conventional DP is *real-time dynamic programming* (RTDP) (Barto *et al.*, 1994). Of the three properties listed above, it has only property 1. RTDP is the result of using sample trajectories—generated by simulation or by the actual control process—to determine the states to which value iteration backups (2) are applied. The backups are applied along these trajectories. This can be thought of as executing asynchronous value iteration (asynchronous because it does not use systematic sweeps) concurrently with a real or simulated control process. The sample trajectory determines the states to which the backup operator is applied, and the current estimate of V^* determines the selection of control actions. The simplest case applies the backup operator only to the states visited in the sample trajectories, but it can be useful to apply the backup operator to neighboring states as well. Instead of performing exhaustive sweeps, RTDP thus attempts to focus its computation on regions of the state set that are likely to be relevant in actual control.

RTDP is most useful when control is constrained to begin in a designated set of initial states. In such cases, an optimal policy need only be defined for states reachable under optimal control from the initial set. Of course, this region is unknown without knowledge of an optimal policy. RTDP can be understood as an attempt to successively confine value iteration to this region. Barto *et al.* (1994) give conditions under which RTDP is guaranteed to converge to a policy whose restriction to this region is optimal in stochastic shortest path problems.

Note that even when applied during actual control, RTDP requires a model of the MDP since it uses the value iteration backup (2). An adaptive version of RTDP (Barto *et al.*, 1994) applicable to the incomplete information case is analogous to the conventional indirect adaptive methods described above. It ties to identify the system and uses its latest model to provide the information necessary to perform backups. Unlike conventional indirect methods, however, this version of RTDP does not execute a complete value iteration computation at each time step, and thus can be applied to problems with very large state sets. RTDP is related to the *Learning-Real-Time A** algorithm of Korf (1990), the *Dyna* architecture of Sutton (1990), the *prioritized sweeping* algorithm of Moore and Atkeson (1993), and to the method proposed by Jalali and Ferguson (1989).

3.2. *Simplifying Backups*

An algorithm called *Q-learning* proposed by Watkins (1989) dispenses with the value iteration backup in favor of an operation that 1) usually requires much less computation, and 2) does not require a model of the MDP. Instead of updating estimates of $V^*(x)$, it updates estimates of optimal action-values $Q^*(x,a)$, which give the expected return for executing $a \in A_x$ when observing x and following an optimal policy thereafter. Let Q_k denote the estimate of Q^* at iteration k. The Q-Learning backup for (x,a) is

$$\begin{aligned} Q_{k+1}(x,a) &= (1-\alpha_k)Q_k(x,a) \\ &+ \alpha_k[r + \gamma \max_{b \in A_y} Q_k(y,b)], \end{aligned} \tag{3}$$

where r is a sample from the reward process with mean $R(x,a)$, y is a sample from the state-transition distribution $p^a(x,\cdot)$, and α_k is a relaxation parameter.

Using a lookup-table to store the values $Q_k(x,a)$, if the Q-learning backup is applied to each state-action pair infinitely often, and some additional natural conditions hold, then Q-learning converges to Q^* with probability one. The Q-learning backup is usually applied along sample trajectories in a manner similar to RTDP. In contrast to the value iteration backup used by RTDP, however, the Q-learning backup does not require a model of the MDP because the observed reward and next state can respectively serve as the samples r and y. They can be sampled on-line from reward and state-transition distributions of the actual system, which can can remain unknown. Additionally, the actions that are optimal with respect to an estimate of Q^* can be determined without a model: for state x, such an action is given by $\arg\max_{a \in A_x} Q_k(x,a)$.

Q-learning is a stochastic approximation of value iteration. By maintaining values for state-action pairs instead of just actions, Q-learning effectively performs updates that are unbiased estimates of value iteration backups. The convergence of Q-learning has been proven from several different perspectives. Watkins and Dayan (1992) provided the first complete proof, and more recent proofs are based on extensions of the theory of stochastic approximation (Jaakkola *et al.*, to appear; Tsitsiklis, 1993). When a single action is available for each state ($|A_x| = 1$, for all $x \in X$), Q-learning reduces to TD(0), one of Sutton's *temporal difference* algorithms (1988) which is applicable to the policy evaluation problem. Bradtke (1994) recently showed how to extend Q-learning to a recursive least squares formulation.

3.3. *Function Approximation*

Backup equations such as (2) and (3) easily can be converted into rules for updating a parameterized approximation of V_k or Q_k. For example, Q_k might be represented as $Q_k(x,a) = f(\theta_k, \phi_x, a)$, where θ_k is the parameter vector at iteration k, ϕ_x is a feature vector representating state x, and f is a given real-valued function differentiable with respect to θ_k for all (ϕ_x, a). Then the backup (3) yields the following gradient-descent parameter update rule:

$$\begin{aligned} \theta_{k+1} &= \theta_k + \\ \alpha_k[r &+ \gamma \max_{b \in A_y} Q_k(y,b) - Q_k(x,a)]\nabla_{\theta_k} Q_k(x,a), \end{aligned} \tag{4}$$

where $\nabla_{\theta_k} Q_k(x,a)$ denotes the gradient vector at (ϕ_x, a) of f as a function of θ_k. However, rules like this present the least understood aspect of DP-based reinforcement learning. It is easy to generate examples showing that these rules can be unstable, or can converge to the wrong function, even under conditions that would seem to be very favorable (e.g., (Bradtke, 1994)). On the other hand, there are also examples in which this approach works very well.

A remarkable demonstration of this, as well as other properties of DP-based reinforcement learning algoritms, is provided by Tesauro's backgammon playing program called *TD-Gammon* (Tesauro, 1992; Tesauro, 1994). Starting with little backgammon knowledge beyond the rules of the game, this program has learned to play near the level of the world's strongest grandmasters (and is still improving). Backgammon can be formulated as a finite MDP whose states are possible states of the game (board configurations plus other information) and whose actions are moves of pieces. Transitions are stochastic because on each turn, a player tosses a pair of die to determine a set of possible moves, and the opponent's responding move can be viewed as an additional stochastic factor. The reward is always zero unless the program wins, in which case it is 1. The value function corresponging to a policy, then, assigns to each state the probability that the program will win from that state using the given policy. The optimal value function assigns the winning probability for optimal play.

Given assumptions about the opponent, this MDP can in principle be solved using DP. However, since there are approximately 10^{20} states, and about 400 possible moves at each play, even a single sweep of value iteration is out of the question (it would take more than 1K years using a 1K MIPS processor), as is a single policy evaluation or policy improvement step of policy iteration. Instead, TD-Gammon played repeated games against an opponent (which happened to be itself). The games were simulated, which is possible because this is an MDP with complete information. A game is a sample trajectory of the MDP.

TD–Gammon maintains and updates an estimate of V^*, which it represents using a multi–layer neural network whose input vector consists of features describing the state of the game, some of which are defined using expert knowledge of backgammon. At each play, it selects the action (move) that leads to the most promising successor state as evaluated by its current estimate of V^*. Then using the estimated value of the actual next state, it updates the network parameters (its weights) by error–backpropagation (Rumelhart *et al.*, 1986) based on a gradient–descent TD rule similar to (4). High–level play was obtained after about 500K games, which took on the order of months of computer time. While this is a significant amount of computation, it is inconsequential compared to what a conventional DP method would require.

4. CONCLUSION

Although the reasons for TD–Gammon's success are still somewhat mysterious, this system illustrates the potential of DP–based reinforcement learning methods for finding useful approximate solutions to large–scale stochastic optimal control problems. Because computation is guided by simulated games, TD–Gammon takes advantage of the fact that very many states have a very low probability of occurring in real games. In this respect, it is similar to other Monte Carlo methods in automatically allocating computation in proportion to that computation's influence on the desired result (Barto and Duff, 1994). TD–Gammon also illustrates that DP–based reinforcement learning can require relatively little computation per–time–step of real or simulated control. The methods that use backups based on reward and next–state samples, instead of the full backups of conventional DP, share the property of stochastic approximation methods in being applicable on–line during actual control when there is no model of the system. Finally, parameterized function approximation, although not well–understood at present, can be used to produce algorithms whose space complexity is within the realm of practicality. Current research is seeking to exploit these properties in a range of large–scale problems.

5. ACKNOWLEDGEMENTS

This research was supported in part by grants from the National Science Foundation (ECS-9214866) and the Air Force Office of Scientific Research, Bolling AFB (AFOSR F49620-93-1-0269).

6. REFERENCES

Barto, A. and M. Duff (1994). Monte Carlo matrix inversion and reinforcement learning. In: *Advances in Neural Information Processing Systems 6* (J. D. Cowan, G. Tesauro and J. Alspector, Eds.). Morgan Kaufmann. San Francisco. pp. 687–694.

Barto, A. G., S. J. Bradtke and S. P. Singh (1994). Learning to act using real-time dynamic programming. *Artificial Intelligence*. In press.

Barto, A.G. (1992). Reinforcement learning and adaptive critic methods. In: *Handbook of Intelligent Control: Neural, Fuzzy, and Adaptive Approaches* (D. A. White and D. A. Sofge, Eds.). pp. 469–491. Van Nostrand Reinhold. New York.

Bertsekas, D. P. (1982). Distributed dynamic programming. *IEEE Transactions on Automatic Control* **27**, 610–616.

Bertsekas, D. P. (1983). Distributed asynchronous computation of fixed points. *Mathematical Programming* **27**, 107–120.

Bertsekas, D. P. (1987). *Dynamic Programming: Deterministic and Stochastic Models*. Prentice-Hall. Englewood Cliffs, NJ.

Bradtke, S. J. (1994). Incremental Dynamic Programming for On-Line Adaptive Optimal Control. PhD thesis. University of Massachusetts, Computer Science Dept. Technical Report 94-62.

Jaakkola, T., M. I. Jordan and S. P. Singh (to appear). On the convergence of stochastic iterative dynamic programming algorithms. *Neural Computation*.

Jacobson, D. H. and D. Q. Mayne (1970). *Differential Dynamic Programming*. Elsevier. New York.

Jalali, A. and M. Ferguson (1989). Computationally efficient adaptive control algorithms for markov chains. In: *Proceedings of the 28th Conference on Decision and Control*. Tampa, Florida. pp. 1283–1288.

Kimble, G. A. (1961). *Hilgard and Marquis' Contitioning and Learning*. Appleton-Century-Crofts. New York.

Korf, R. E. (1990). Real-time heuristic search. *Artificial Intelligence* **42**, 189–211.

Kumar, P. R. (1985). A survey of some results in stochastic adaptive control. *SIAM Journal of Control and Optimization* **23**, 329–380.

Kumar, P. R. and P. Varaiya (1986). *Stochastic Systems: Estimation, Identification, and Adaptive Control*. Prentice-Hall. Englewood Cliffs, NJ.

Mendel, J. M. and R. W. McLaren (1970). Reinforcement learning control and pattern recognition systems. In: *Adaptive, Learning and Pattern Recognition Systems: Theory and Applications* (J. M. Mendel and K. S. Fu, Eds.). pp. 287–318. Academic Press. New York.

Minsky, M. L. (1961). Steps toward artificial intelligence. *Proceedings of the Institute of Radio Engineers* **49**, 8–30. Reprinted in E. A. Feigenbaum and J. Feldman, editors, *Computers and Thought*. McGraw-Hill, New York, 406–450, 1963.

Moore, A. W. and C. G. Atkeson (1993). Prioritized sweeping: Reinforcement learning with less data and less time. *Machine Learning* **13**(1), 103–130.

Norman, J. M. (1972). *Heuristic Procedures in Dynamic Programming*. Manchester University Press. Manchester, UK.

Ross, S. (1983). *Introduction to Stochastic Dynamic Programming*. Academic Press. New York.

Rumelhart, D. E., G. E. Hinton and R. J. Williams (1986). Learning internal representations by error propagation. In: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, vol.1: Foundations* (D. E. Rumelhart and J. L. McClelland, Eds.). Bradford Books/MIT Press. Cambridge, MA.

Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal on Research and Development* pp. 210–229. Reprinted in E. A. Feigenbaum and J. Feldman, editors, *Computers and Thought*, McGraw-Hill, New York, 1963.

Samuel, A. L. (1967). Some studies in machine learning using the game of checkers. II—Recent progress. *IBM Journal on Research and Development* pp. 601–617.

Sutton, R. S. (1988). Learning to predict by the method of temporal differences. *Machine Learning* **3**, 9–44.

Sutton, R. S. (1990). Integrated architectures for learning, planning, and reacting based on approximating dynamic programming. In: *Proceedings of the Seventh International Conference on Machine Learning*. Morgan Kaufmann. San Mateo, CA. pp. 216–224.

Sutton, R. S., Ed.) (1992). *A Special Issue of Machine Learning on Reinforcement Learning*. Vol. 8. *Machine Learning*. Also published as *Reinforcement Learning*, Kluwer Academic Press, Boston, MA 1992.

Tesauro, G. J. (1992). Practical issues in temporal difference learning. *Machine Learning* **8**, 257–277.

Tesauro, G. J. (1994). Temporal difference learning in TD-Gammon. *Communications of the ACM*. in press.

Tsitsiklis, J. N. (1993). Asynchronous stochastic approximation and Q-learning. Technical Report LIDS-P-2172. Laboratory for Information and Decision Systems, MIT. Cambridge, MA.

Watkins, C. J. C. H. (1989). Learning from Delayed Rewards. PhD thesis. Cambridge University. Cambridge, England.

Watkins, C. J. C. H. and P. Dayan (1992). Q-learning. *Machine Learning* **8**, 279–292.

Werbos, P. J. (1977). Advanced forecasting methods for global crisis warning and models of intelligence. *General Systems Yearbook* **22**, 25–38.

Werbos, P. J. (1992). Approximate dynamic programming for real-time control and neural modeling. In: *Handbook of Intelligent Control: Neural, Fuzzy, and Adaptive Approaches* (D. A. White and D. A. Sofge, Eds.). pp. 493–525. Van Nostrand Reinhold. New York.

THE MODULAR ORGANIZATION OF MOTOR CONTROL: WHAT FROGS CAN TEACH US ABOUT ADAPTIVE LEARNING

F.A. MUSSA–IVALDI* and E. BIZZI**

**Northwestern University Medical School, Departments of Physiology, Physiscal Medicine and Rehabilitation, Chicago, Ill, USA*

***Massachusetts Institute of Technology, Department of Brain and Cognitive Sciences, Cambridge, MA, USA*

Abstract. Investigations of motor control in the spinaized frog have suggested the presence, within the spinal cord, of a number of discrete control modules that may act, either individually or in combination, to generate a variety of motor behaviors. The experimental findings indicate that the simultaneous activation of two modules leads to the vectorial summation of their output forces. These findings have led us to consider the output generated by the spinal cord as a system of vector–valued basis functions. In this context, the execution of complex motor behaviors and the adaptation to novel environments may be investigated as forms of function approximation.

Key Words. Motor control, Bio control, Function approximation, Impedance control, Adaptation

1. INTRODUCTION

Recent evidence obtained by focal microstimulation of the frog's spinal cord has revealed that the lumbar grey matter contains a number of circuits that are organized to coordinate the activation of multiple muscles (Bizzi et al. 1991, Mussa-Ivaldi and Giszter, 1992). As a consequence, the mechanical effect of activating a small region in the lumbar grey matter is a pattern of forces that may be measured at the ankle and that converge toward an equilibrium point in space. In this article we explore some of the implications of these findings with respect to the task of generating a variety of motor behaviors and of adapting the learned behaviors to changing mechanical environments.

2. THE MODULAR ORGANIZATION OF THE FROG'S SPINAL CORD

We begin by discussing some recent experiments which were directed at identifying the organization of motor output in the frog's spinal cord. (Bizzi *et al.*, 1991; Giszter *et al.*, 1993). In these experiments, Bizzi, Giszter and Mussa–Ivaldi stimulated with a microelectrode the gray matter of the lumbar cord and measured isometric forces at the ankle with the frog's leg held in a variety of postures. Each stimulus consisted in a train of current impulses which lasted typically 300 milliseconds. The impulses had a duration of one millisecond and a frequency of 40 Hz. The peak amplitude ranged between 1 and 6 microamperes.

The analysis of the mechanical responses was restricted to the horizontal plane. Ankle locations were specified as pairs of Cartesian coordinates with respect to the hip joint. The ankle was placed at a number of locations (9 to 16), which formed a regular grid over the leg's workspace. At each location, the force vector, F, elicited by the stimulation of the same spinal-cord locus and with the same current parameters was recorded. The measured force vector varied as a function of the workspace location, x . The changes in force through the workspace may be attributed to a number of factors: the neural activation of muscles, as well as their lengths, moment arms and visco-elastic properties.

The data collected in this microstimulation experiment constitute a set of samples of a time-varying force field, $F(x,t)$, obtained at a number of tested sites. F(x, t) denotes the force vector at the location, r, after a latency, t, from the onset of the stimulus. In most instances, the spatial distribution of the forces induced by the stimulation was structured in a well defined pattern: the flow of force vectors converged towards a single equilibrium point. The area of the spinal cord from which such convergent fields were observed was quite broad. Force fields converging to different equilibrium points have been observed by placing the stimulating electrode in different sites of the spinal cord. These experiments indicated that there are at least four spinal regions from which four distinct convergent fields are elicited. Within each region, the stimulation of the pre-motor layers of the spinal gray matter generates a similar pattern of force vectors.

In all likelihood, the convergent force fields derive at least in part from the activation of the spinal interneurons. The branches of these interneurons make synaptic connections with different motoneurons[1] Remarkably, when the electrode was placed within a motoneuron pool, the stimulation induced a radically different type of field. The force vectors were arranged in a parallel or divergent pattern instead of converging to an equilibrium point. This result indicates that the convergence of the field is likely due to a critical balance of different muscles. Balance that is lost when the electrical stimulus activates only the motoneurons connected to a single muscle or to a group of synergistic muscles.

In summary, these microstimulation experiments suggested that the interneurons in the spinal cord may be organized in a relatively small number of modules. Each module is connected to a balanced set of motoneuron pools. Thus, each module organizes the mechanical behavior of a group of spring-like muscles into an end-point field that converges to a single equilibrium position. We are led then to ask whether and how the central nervous system may generate not just a few fields but a full repertoire of them by combining the outputs of a few modules. This question has been addressed by stimulating two separate spinal sites, first independently and then simultaneously (Bizzi *et al.*, 1991; Mussa-Ivaldi *et al.*, 1994). Remarkably, in the large majority of the cases (87.8%) the simultaneous stimulation of two sites resulted in a force field proportional to the vector sum of the fields obtained by the stimulation delivered at each point. This finding indicates that the central nervous system may generate a repertoire of fields by a simple superposition mechanism.

3. ENDPOINT SUMMATION AND KINEMATIC REDUNDANCY

The finding of vector summation has been particularly surprising when the frog's leg was connected to the force sensor in a *kinematically redundant* configuration (Mussa-Ivaldi *et al.*, 1994). In this case, the foot was attached to the sensor through a gimbal mechanism which allowed the leg to assume a variety of configurations while the center of the foot was held rigidly at a location in the plane. Theoretical considerations are sufficient to establish that the endpoint forces of a kinematically redundant limb need not to combine according to vector summation, even in the simple case of direct muscle activation (Gandolfo and Mussa-Ivaldi, 1993).

The redundancy of a limb is established by the imbalance between the limb's degrees of freedom and the number of end-point coordinates. For example, in our experiment the configuration of the frog's leg was specified by six joint angles: three at the hip, one at the knee and two at the ankle. We indicate these angles collectively by the array $q = (q_1, q_2, \ldots, q_6)$. At the same time, the center of the foot was constrained by the gimbal to remain on the horizontal plane. Thus, the position of the foot was determined by a two-dimensional array: $x = (x_1, x_2)$. With such an imbalance between degrees of freedom and end point coordinates, the mapping from q to x cannot be inverted to derive uniquely the leg configuration corresponding to any given foot location.

The generalized force in the leg's configuration space is a 6-dimensional torque vector, $\tau = (\tau_1, \tau_2, \ldots, \tau_6)$. The activation of a viscoelastic muscle, m, is expected to give rise to a field of generalized forces in configuration space, that is a mapping, ϕ, from q to τ: $\tau = \phi(q)$. As far as one is dealing with configuration space– the space of all independent degrees of freedom– and with a set of independent actuators, one can prove that the mapping $\phi(q)$ obeys the rule of vectorial summation (Gandolfo and Mussa-Ivaldi, 1993): given two distinct muscles, m and n, with torque fields $\phi_m(q)$ and $\phi_n(q)$, their simultaneous action gives rise to the field

$$\phi_\Sigma(q) = \phi_m(q) + \phi_n(q). \qquad (1)$$

It is crucial to stress that in the above expression the three fields, ϕ_Σ, ϕ_m and ϕ_n are evaluated at the same configuration, q. In contrast, with three different configurations, q_a, q_b and q_c it is possible that

$$\phi_\Sigma(q_a) \neq \phi_m(q_b) + \phi_n(q_c). \qquad (2)$$

even when Equation (1) is satisfied.

Let us now consider the field of forces generated at the endpoint of the redundant arm by the activation of two independent actuators. Each one of these fields may be regarded as the outcome of an experiment in which a muscle is activated and the end-point force vector $F = (F_1, F_2)$ is measured at a number of locations in the x_1, x_2

[1] Motoneurons constitute the final output stage of the spinal cord. They are organized in clusters, or "pools", of cells which innervate the same muscle.

plane. Let us indicate by $F_m(x)$, $F_n(x)$ and $F_\Sigma(x)$ the endpoint force fields obtained by activating, respectively, muscle m, muscle n and both muscles simultaneously. If one assumes that Equation (1) is true, one can inquire whether $F_\Sigma(x) = F_m(x) + F_n(x)$.

The answer is no. Because of the kinematic redundancy, the same hand location in the above expression will eventually correspond to three different joint configurations, q_m, q_n and q_Σ. Therefore, in spite of the fact that we are looking at the same endpoint location, we are now in the case described by Equation (2) rather than (1)

The above considerations establish that the endpoint fields of a redundant limb need not combine according to vectorial summation, even in the simple case in which these fields are induced by direct muscle stimulation. However, we cannot rule out vector summation in particular cases nor are we able yet to establish how non-linear the behavior of a limb is expected to be given the limb's biomechanical structure. Remarkably, the experimental results of Mussa-Ivaldi et al. (1994) have indicated that endpoint field summation is the predominant outcome in the presence of kinematic redundancy. This finding is important as it suggests that *the neural controller of a redundant mechanical system may neglect the details of the system's kinematics and focus exclusively at the "end point" that is at the site of interaction between the system and its environment.* This is one of the main aspects of what is now known as "impedance control" (Hogan, 1985).

4. FIELD APPROXIMATION

The findings described in the previous sections lead to a formal analysis of motor control based on the framework of function approximation (see also (Sanner and Slotine, 1992)). Let us consider a simplified system consisting of a set of K independent control modules acting in parallel upon the muscles of a multi-joint limb. Each control module establishes the viscoelastic properties of a group of muscles. As a result, each module generates at the interface between the limb and the environment a force-field

$$F_i = \Phi_i(x, u_i). \qquad (3)$$

where u_i indicates the controller's command variable. The findings of the microstimulation experiments suggest that the net force field generated by the whole ensemble of controllers is:

$$F(x, u_1, \ldots, u_K) = \sum_{i=1}^{K} \Phi^i(x, u_i). \qquad (4)$$

We may define the *repertoire* of a control network as the set of all possible net fields with all possible values of the control variables. If we denote by $\mathcal{U}_i$ the set of the admissible values for the command variable u_i, the repertoire of the control networks is the set:

$$X = \left\{ \sum_{i=1}^{K} \Phi^i(x, u_i) \mid u_i \in \mathcal{U}_i \right\}$$

Vector fields provide not only a description of the control processes but also a framework for specifying the planning of a desired behavior. In robotics some researchers have proposed to specify the planning of complex tasks by force fields defined over a manipulator's workspace. For example, Kathib (Kathib, 1986) used potential fields to represent the planning of reaching motions within obstacle-ridden environments. Kathib (1986) and Hogan (1984) suggested to implement the planned potential functions by a set of controllers operating in parallel. According to this view, each controller derives the forces induced by a target or by an obstacle by directly computing the gradient of the corresponding potential field. Thus, the field specified by the planner is faithfully implemented by the controllers.

Here, we present a different approach in which both planning and execution are directly represented as vector fields instead of scalar potential fields. A vectorial representation is indeed more general than a scalar representation. It is always possible to derive the former from the latter, whereas a vector field may not be reducible to a potential function. Second, we may reasonably assume that a vertebrate's motor control system has a repertoire, X, generated by the vector combination of independent modules. This repertoire is limited by the pre-defined properties of the controllers and of the actuators. In other words we cannot assume that any goal may be accurately mapped into a corresponding motor output. Instead, the spinal microstimulation experiments suggest that control primitives are established a-priori by the connections between muscles and neural circuits. Following this view, the execution of an arbitrary motor plan is equivalent to a field-approximation problem: given a planned field, $P(x)$, the problem is to find a field, $F(x) \in X$, which minimizes some norm

$$\| P(x) - F(x) \|^2$$

suitably defined over the limb's state space (see also the section on motor control in Poggio, 1990).

The field-approximation problem assumes a more tractable form if a planned behavior is specified by a finite set of M force vectors, $\{P^1, P^2, \ldots, P^M\}$, defined at M points, $\{x^1, x^2, \ldots, x^M\}$, rather than by a continuous field, $P(x)$. We consider these vectors as samples (or "examples") of a field to be filled in (or "completed") by an appropriate choice of control parameters.

For the sake of clarity we may further simplify our discussion by using a factored form for the the output function (3) implemented by each control module:

$$\Phi_i(x, u_i) = u_i \phi^i(x). \quad (5)$$

We call this form *linear tuning*: the control variable, u_i plays the role of a scaling factor applied to a (nonlinear) output field. In this case, the repertoire of the control network is the linear span:

$$X = \left\{ \sum_{i=1}^{K} u_i \phi^i(x) \mid u_i \in \mathcal{U}_i \right\}$$

and we may apply the methods of linear algebra as outlined in Mussa-Ivaldi (1992) for approximating an arbitrary pattern of vectors. Given a set of M pattern vectors, $P^1, \ldots, P^M$, specified at M distinct locations, $x^1, \ldots, x^M$, the approximation goal is to find a set of control parameters such that

$$\sum_{i=1}^{K} u_i \phi^i(x^j) \sim P^j$$

for $j = 1, \ldots, M$. In an N-dimensional vector space, this goal is expanded into a system of MN linear equations (one per data component) in K unknowns. This system of equations can be compactly written as:

$$\Phi u = \hat{P} \quad (6)$$

where we have introduced the (unknown) control vector

$$u = (u_1, u_2, \ldots, u_K),$$

the matrix

$$\begin{bmatrix} \phi_1^1(x^1) & \phi_1^2(x^1) & \ldots & \phi_1^K(x^1) \\ \phi_1^1(x^2) & \phi_1^2(x^2) & \ldots & \phi_1^K(x^2) \\ & \ldots & & \\ \phi_1^1(x^M) & \phi_1^2(x^M) & \ldots & \phi_1^K(x^M) \\ & \ldots & & \\ & \ldots & & \\ \phi_N^1(x^1) & \phi_N^2(x^1) & \ldots & \phi_N^K(x^1) \\ \phi_N^1(x^2) & \phi_N^2(x^2) & \ldots & \phi_N^K(x^2) \\ & \ldots & & \\ \phi_N^1(x^M) & \phi_N^2(x^M) & \ldots & \phi_N^K(x^M) \end{bmatrix},$$

and the "pattern vector"

$$\hat{P} = (P_1^1, P_1^2, \ldots, P_1^M, P_2^1, P_2^2, \ldots, P_2^M, \ldots, P_N^1, P_N^2, \ldots, P_N^M).$$

The approximation problem (6) can be solved for any set of pattern vectors if the output fields, $\phi^i(x)$, are linearly independent. In this case, the output fields– which we call *basis fields* – form a basis for a K-dimensional functional space. It is possible to derive basis fields from scalar basis functions as detailed in (Mussa-Ivaldi, 1992)[2]. Following this approach, a general representation of a continuous field is obtained by combining K_I irrotational basis fields and K_S solenoidal basis fields:

$$F(x) = \sum_{i=1}^{K_I} c_i \varphi^i(x) + \sum_{i=1}^{K_S} d_i \psi^i(x).$$

An irrotational basis field, $\varphi^i(x)$, has zero curl and is the gradient of some scalar basis function, $g_i(x)$, that is

$$\varphi^i(x) = \nabla g_i(x) \quad \text{and} \quad \text{curl}(\varphi^i) = 0$$

A solenoidal basis field, $\psi^i(x)$, has zero divergence. In the Euclidean metric, an irrotational basis field is obtained from a basis function, $g_i(x)$, by applying the operator $A\nabla$, where A is an antisymmetric[3] matrix:

$$\psi^i(x) = A\nabla g_i(x) \quad \text{and} \quad \text{div}(\psi^i) = 0.$$

This statement can be demonstrated for $x \in \Re^N$, with $N \geq 2$. In particular, with N = 2 and N = 3 an antisymmetric matrix, A, corresponds to a 90^o rotation operator.

The set of control coefficients, c_i and d_i, obtained from the approximation of a field feature can be regarded as a *representation* of the corresponding planning goal within the repertoire of the controller network.

5. ADAPTIVE CONTROL

The presence of a set of independent control modules provides a powerful framework for the representation of motor learning. We may distinguish between two broad classes of motor learning: *skill learning* and *adaptive learning*. The first type of learning refers to the acquisition of new patterns of control that lead to the execution of novel behaviors. This happens for example, when we learn

[2] See also (Wahba, 1982) for a similar generalization of splines to the approximation of vector fields on the sphere.

[3] A matrix, $A = [a_{i,j}]$ is said to be antisymmetric if $a_{i,j} = -a_{j,i}$ $(a_{i,i} = 0)$.

to play tennis, or to drive a car. In contrast, adaptive learning is the process by which a repertoire of previously acquired skills is transported across different environments. To put it in more formal terms, consider the following differential equation that describes the behavior of a broad class of dynamical systems:

$$D(x, \dot{x}, \ddot{x}) = C(x, \dot{x}, t) \quad (7)$$

The term, $D(\cdot)$ on the left—hand side is a non-linear time–independent second–order differential expression that captures the dynamics of a plant at the interface with the environment. For example, the controlled system may be an arm in which case x indicates the position of the hand and $D(\cdot)$ is the force generated at the hand in response to an externally imposed displacement. The term $C(\cdot)$ on the right–hand side of (7) is the time– and state–dependent force generated by a controller upon the system described by $D(\cdot)$. Equation (7) is a second–order ordinary differential equation which defines a unique trajectory, $x(t)$ for any given initial condition, $x_0 = x(t_0)$, $\dot{x}_0 = \dot{x}(t_0)$. In this context, learning new skills corresponds to establishing a repertoire of controllers, $C_i(x, \dot{x}, t)$, that generate a corresponding repertoire of trajectories, $x_i(t)$. The task of adaptive learning can be instead described as follows. Suppose that a new load is attached to the controlled system. This load may be represented as an additional dynamical term, $E(x, \dot{x}, \ddot{x})$. With the controller of Equation (7), the differential equation describing the system with the load is:

$$D(x, \dot{x}, \ddot{x}) + E(x, \dot{x}, \ddot{x}) = C(x, \dot{x}, t). \quad (8)$$

Generally, this equation admits a new solution, $x^*(t)$, given the same initial conditions, $(x_0, \dot{x}_0)$. The purpose of adaptive learning is to generate a correction of the old controller so as to restore the behavior of the system as it was generated by Equation (7) before the onset of the load. A possible way to represent this correction is by means of a new controller, C', expressed as the old one, C, plus an additive term, D:

$$C'(x, \dot{x}, t) = C(x, \dot{x}, t) + D(x, \dot{x}, t).$$

In this case, it is not difficult to see that a sufficient condition for restoring the old controller is that, along the original trajectory $x(t)$, the following differential equation be satisfied:

$$E(x, \dot{x}, \ddot{x}) = D(x, \dot{x}, t). \quad (9)$$

If this is the case, the equation

$$D(x, \dot{x}, \ddot{x}) + E(x, \dot{x}, \ddot{x}) = C'(x, \dot{x}, t)$$

is reduced to the original equation (7) for the system without load. We must stress however, that this approach to load compensation is by no means unique. In fact, a strategy based upon the increase of the controller's impedance about the desired trajectory, $x(t)$, would be equally adequate to restore the original behavior (Shadmehr and Mussa–Ivaldi, 1994).

The approach described by Equation (9) is equivalent to state that adaptation is achieved by constructing an internal model (the controller field D) of the environment (the load field, E). Recent experiments of Shadmehr and Mussa–Ivaldi (1994) have indeed suggested that the construction of such an internal model may be the process through which human subjects achieve the adaptation of multi–joint arm movement to external loads. In these experiments, subjects were asked to execute reaching movements of the hand in the horizontal plane while holding the handle of a robot manipulandum. During the initial part of each experiment, the manipulandum acted as a low–impedance passive device (the motors were turned off) and the subjects were asked to move their hand toward a set of targets located at different location of the horizontal plane. Under these circumstances, the hand movements were characterized by smooth straight–line hand trajectories, as described by several authors (Morasso, 1981; Flash and Hogan, 1985). In a second part of the experiment, the robot manipulandum was programmed to generate a force field, that is a function that mapped the instantaneous hand velocity into a corresponding force applied to the subject. Since the force field significantly changed the dynamics of the task, subjects' initial movements in the force field were grossly distorted as compared to their movements in free space. However, with practice, hand trajectories in the force field converged to a smooth and straight path, very similar to that observed in free space. In order to investigate the mechanism underlying this adaptation, Shadmehr and Mussa–Ivaldi (1994) observed the response to sudden removal of the field after a training phase. The resulting trajectories were approximately mirror images of those which were observed when the subjects were initially exposed to the field. This suggested that the arm controller had indeed composed an internal model of the applied field. A model that, like the field D in Equation (9) the nervous system used to predict and compensate the forces imposed by the environment.

The construction of an internal model is quite consistent with the field–approximation framework described in the previous section (Sanner and Slotine, 1992). Indeed, let us suppose that a set of modules or "pattern generators" is implemented– for example– by the neural circuits within a subject's spinal cord and let us suppose that each pattern generator when active induces a field $\phi_i(x, \dot{x}, t|c_i)$ of endpoint forces. In this expression, the variable, c_i is an "activation variable" representing a tuning effect of descending commands. Using this representation, the adaptive controller described by Equation (9) can be written as

$$E(x, \dot{x}, y(t)) \sim \sum_i \phi_i(x, \dot{x}, t|c_i) \quad (10)$$

where $y(t) = \ddot{x}(t)$ and the sum is extended over a particular set of spinal modules. The above expression describes a field approximation problem whose goal is to derive a set of control variables, c_i, that minimize some norm

$$\| E - \sum_i \phi_i \|^2$$

defined over the space of endpoint trajectories.

6. CONCLUSION

The microstimulation of the frog's spinal cord has provided us with important evidence suggesting that the output stages of the vertebrate motor system is organized into a small set of independent control elements. Each control element coordinates the activations of a group of viscoelastic muscles and the observable consequence of these activations is a field of forces acting upon the controlled limb.

The modular nature of these control elements is expressed by the observation that their simultaneous activation leads to the vectorial summation of the corresponding fields. Remarkably, this summation property is preserved when the forces are measured at the endpoint of a kinematically redundant limb. We are led by these findings to conclude that the force fields induced by the muscle spring–like properties at the interface between a controlled limb and its environment may constitute a "vocabulary" of simple control functions that are combined to construct more complex behaviors. From this standpoint, our findings are consistent with the general framework of impedance control (Hogan, 1985).

Acknowledgments: This research was supported by NIH grants NS09343 and AR26710 and ONR grant N00014/90/J/1946.

7. REFERENCES

Bizzi, E., F. A. Mussa-Ivaldi and S. F. Giszter (1991). Computations underlying the execution of movement: A biological perspective. *Science* **253**, 287 – 291.

Flash, T. and N. Hogan (1985). The coordination of arm movements: An experimentally confirmed mathematical model. *Journal of Neuroscience* **5**, 1688 – 1703.

Gandolfo, F and F. A. Mussa-Ivaldi (1993). Vector summation of end–point impedance in kinematically redundant manipulators. In: *Proceedings of th3 1993 IEEE/RSJ International Conference on Intelligent Robots and Systems.* Yokohama, Japan.

Giszter, S.F., F. A. Mussa-Ivaldi and E. Bizzi (1993). Convergent force fields organized in the frog's spinal cord. *Journal of Neuroscience* **13**, 467–491.

Hogan, N. (1984). Some computational problems simplified by impedance control. In: *ASME Computers in Engineering.* Las Vegas.

Hogan, N. (1985). Impedance control: an approach to manipulation: Parts i, ii, iii. *ASME Journal of Dynamic Systems, Measurement and Control* **107**, 1 – 24.

Kathib, O. (1986). Real-time obstacle avoidance for manipulators and mobile robots. *The International Journal of Robotics Research* **5**, 90 – 99.

Morasso, P. (1981). Spatial control of arm movements. *Experimental Brain Research* **42**, 223–227.

Mussa-Ivaldi, F. A. (1992). From basis functions to basis fields: Using vector primitives to capture vector patterns.. *Biological Cybernetics* **67**, 479–489.

Mussa-Ivaldi, F. A. and S. F. Giszter (1992). Vector field approximation: A computational paradigm for motor control and learning. *Biological Cybernetics* **67**, 491–500.

Mussa-Ivaldi, F. A., S. F. Giszter and E. Bizzi (1994). Linear combinations of primitives in vertebrate motor control. *Proceedings of the National Academy of Sciences USA* **91**, 7534 – 7538.

Poggio, T. (1990). A theory of how the brain might work. *Cold Spring Harbor Symposium on Quantitative Biology: The Brain* **55**, 899 – 910.

Sanner, R. M. and J-J. E. Slotine (1992). Gaussian networks for direct adaptive control. *IEEE Transactions on Neural Networks* **3**, 837 – 863.

Shadmehr, R. and F. A. Mussa-Ivaldi (1994). Adaptive representation of dynamics during learning of a motor task. *Journal of Neuroscience* **14**, 3208 – 3224.

Wahba, G. (1982). Vector splines on the sphere, with application to the estimation of vorticity and divergence from discrete noisy data. In: *Multivariate Approximation Theory* (W. Schempp and K. Zeller, Eds.). pp. 407 – 429. Birkhauser Verlag.

TRAJECTORY LEARNING AND CONTROL MODELS: FROM HUMAN TO ROBOTIC ARMS

T. FLASH*

*Weizmann Institute of Science, Department of Applied Mathematics and Computer Science, Rehovot 76100, Israel

Abstract. Based upon combined experimental and modeling studies, in recent years computational models of different aspects of human arm trajectory control have been developed. These models have dealt with the topics of human arm trajectory planning, on-line trajectory modification, the control of contact tasks and motor learning and adaptation. In this paper we discuss several of these models and the corresponding motion control algorithms for robotic manipulators which were developed based on the insights gained from our human motor control studies.

Key Words. Motor control; robotics; robot manipulators; trajectory planning: adaptation; neural network models

1. INTRODUCTION

In principle, research into biological and artificial motion control is complementary. Biological motor control studies are aimed at investigating motor organization through the analysis of human movements. Robotics research, on the other hand, is aimed at developing intelligent control systems capable of controlling the motion of man-made robots and manipulators. Although natural and artificial systems seem to employ very different computational machineries and vastly different actuators and sensors, both systems share the same goal of connecting perception with action and at higher levels of abstraction must cope with similar problems.

Robotics, therefore, can serve as a useful medium for suggesting and testing ideas about the control of biological movement since it allows to test whether these ideas can be made to work in practice. On the other hand, since human motor control capabilities far surpass those of artificial systems, human motor control studies can provide us with new ideas of how to design and control robotic systems. In this paper we review some of our human motor control studies in which we have focused on arm trajectory control with a special emphasis on the topics of hand trajectory planning and modification, motor execution and motor adaptation and learning. These are also topics of fundamental importance in robotics research. Hence, in this paper we discuss the main findings and models that emerged from our human studies and the corresponding robot control algorithms that were developed, based on these models.

2. TRAJECTORY PLANNING AND MODIFICATION

2.1. *Human Arm Trajectory Planning*

A research topic that has attracted great attention among human motor control scientists over the last decade is arm trajectory formation. Since motor behavior is fundamentally multi-dimensional, movements might be alternatively represented in muscle, joint or end-effector spaces. Hence, a fundamental question in motor control research is in what space(s) or coordinate frame(s) does the brain represent movement. Another fundamental question is what rules or principles govern the selection of specific movement trajectories among the infinite number of possible ones. Experimental observations of human unconstrained point-to-point movements have revealed the tendency of human subjects to generate roughly straight hand paths with unimodal bell-shaped speed profiles (Morasso, 1981; Flash and Hogan, 1985). Given that the common invariant features of those movements were only evident in the trajectories of the hand through external space and not in the movements of individual limb segments, those results have provided strong indication that planning takes place in terms of hand trajectories rather than in terms of joint rotations. In curved movements, although the hand paths appeared smooth, movement curvature was not uniform and the hand trajectories typically displayed two or more curvature maxima. The hand velocity profiles also had two or more peaks and the minima between adjacent peaks corresponded to the maxima in curvature (Flash and Hogan, 1985).

To account for those experimental observations, a model was developed suggesting that a major objective of motor coordination is to achieve the smoothest possible hand trajectories under the circumstances (Flash and Hogan, 1985). Equating maximum smoothness with the minimization of hand jerk (the rate of change of hand acceleration), the unique trajectories which yield the best performance were mathematically determined and compared to the measured movements. This model has succeeded in accounting for both the qualitative features and the quantitative details of the observed behavior. Moreover, since the minimum-jerk criterion function was expressed in terms of Cartesian hand coordinates, the success of the model in capturing the essential kinematic characteristics of real movements has provided strong evidence in support of the idea that arm movements are internally represented in terms of the movement of the hand through external space.

2.2. *Human Arm Trajectory Modification*

Beyond the need to generate movements toward static objects, the nervous system must also correct or modify ongoing movements and control dynamic tasks. Thus, in a recent series of studies the mechanisms subserving arm trajectory modification were investigated in human subjects, using the double-step target displacement paradigm. In this paradigm, the subject is presented with a visual target and is instructed to move his/her arm toward the target. In some of the trials, however, the target might be suddenly displaced following an inter-stimulus-interval (ISI) to a new location either during the reaction or movement time (Flash and Henis, 1990) and the arm trajectory must be accordingly modified. In considering the problem of rapid modification of an ongoing motion plan, one's first guess is that upon the change in target location the initial motion plan is aborted and the motor system plans a new movement toward the new target location (this will be referred to here as the "abort-replan" scheme). However, mathematical analysis of experimentally recorded human arm movements has demonstrated that the kinematic features of the measured movements were better accounted for by an alternative trajectory modification scheme (Flash and Henis, 1990). According to this scheme, the initial trajectory plan is neither aborted nor modified following the target switch but is instead vectorially summed with a new unconstrained point-to-point trajectory plan for moving between the first and second target locations. The second added trajectory plan was also shown to have similar kinematic characteristics to those of unconstrained point-to-point trajectories, i.e., to follow a kinematic profile which is adequately described by the minimum-jerk model. Thus, the modified trajectories appear to result from the vectorial superposition of two time-shifted point-to-point motions. This scheme, is computationally simpler than the alternative abort-replan scheme, since by contrast to the abort-replan scheme, no information is required about the expected hand location at the switching time. Moreover, the superposition scheme was found to be significantly more successful in accounting for the kinematic features of the measured modified movements than the alternative abort-replan scheme. These findings have indicated the possibility that arm trajectory modification involves parallel planning and superposition of elementary motion units and that there exist a basic repertoire of elementary movements from which more complex movements are constructed.

More recently, arm trajectory modification was further investigated for relatively short ISIs, i.e., ISIs ranging between 10-200 msecs (Henis and Flash, 1994). Under such conditions, a high percentage of the movements were found to be initially directed in between the first and second target locations (*averaged* trajectories). The initial direction of motion was found to depend on D: the time difference between the presentation of the second stimulus and movement onset. In attempting to account for the kinematic features of the averaged trajectories, the abort-replan and the superposition schemes were again used, and the performance of the superposition scheme was found to surpass that of the abort-replan scheme. This time, however, the modified trajectories were shown to result from the vectorial summation of the two following elemental motions: one for moving between the initial hand position and an intermediate location, and a second one for moving between the intermediate location and the final target. Moreover, it was hypothesized that due to the quick displacement of the stimulus, the *internally specified* intermediate goal might be influenced by both stimuli. Hence, its location was hypothesized to be different from that of the first stimulus. Mathematical analysis was further performed to infer the intermediate (i.e., internally represented) target locations. This analysis has shown that for increasing values of the D parameter, these inferred locations gradually shift from the first toward the second target locations along a path that curved toward the initial hand position. These locations showed also a strong resemblance to the intermediate locations of saccadic eye movements generated in a similar double-step paradigm. Such similarities in the specification of target locations used in the generation of eye and hand movements may serve to simplify the computations and coordinate transformations underlying eye-hand coordination and visuomotor integration (Henis and Flash, 1994).

2.3. *Motion Planning Algorithms for Robot Manipulators*

The aforementioned studies have inspired the development of motion planning algorithms for robot manipulators. For example, maximizing the trajectory smoothness, i.e., minimizing hand jerk, was one of the objectives of a computationally efficient trajectory planner developed by Flash and Potts (1988). In another study (Flash, 1990), efficient trajectory planners, based on the use of multi-grid solvers, were developed in order to derive the optimal trajectories for a multi-joint manipulator in the case of highly nonlinear objective functions such as the minimization of the rate of change of joint torques. This objective function is closely related to the minimization of hand jerk but gives different predictions in the case of multi-joint movements (see below). Furthermore, following our human arm trajectory modification studies, a trajectory modification scheme for robot manipulators based on the superposition scheme was developed (Gat-Falik, 1990). The purpose of that work was to deal with situations whereby an unexpected target displacement raises the need for mid-flight modifications of the manipulator's ongoing motions. These algorithms, however, might also prove to be useful when a manipulator is moving toward objects, placed on a conveyer belt, or during the performance of other dynamic tasks, for example when a new or more accurate information about target location becomes available in the course of the movement. One advantage of the proposed superposition scheme over other more standard modification schemes is that it does not require knowledge about the kinematic state of the end-effector at the switching time. Thus, the basic superimposed trajectory plans can be carried out in parallel while giving rise to smooth transitions between consecutive movements. This may permit using a multi-processor system with relatively little synchronization and communication. The basic trajectories used were again assumed to be minimum-jerk quintic polynomials in time as follows:

$$x(t) = (x_1 - x_0)\left(10\left(\frac{t}{T_1}\right)^3 - 15\left(\frac{t}{T_1}\right)^4 + 6\left(\frac{t}{T_1}\right)^5\right) \quad (1)$$

where x_0 and x_1 are the initial end-effector and target positions, respectively, t is time and T_1 is the first unit duration. The form of the added second trajectory was assumed to be:

$$x(t) = (x_2 - x_1)\left(10\left(\frac{t-s}{T_2}\right)^3 - 15\left(\frac{t-s}{T_2}\right)^4 + 6\left(\frac{t-s}{T_2}\right)^5\right) \quad (3)$$

where s is the time of the target switch, x_2 is the new target location and T_2 is the duration of the added trajectory unit.

Several problems related to the use of the superposition scheme for robot manipulators were considered (Gat-Falik, 1990). The first problem involved finding the minimum traveling time for the entire motion without violating kinematic constraints such as lower and upper bounds on the end-effector velocities and accelerations. This reduces to the problem of determining the optimal switching time s (following the actual target switch) and the durations of the superimposed units T_1 and T_2. This problem was solved by using nonlinear optimization methods while assuming that the knowledge about the new target location is incremental. Thus, only the set of parameters for the new unit must be determined while the preceding trajectory units are not modified. Position constraints in joint space were also incorporated into the superposition scheme. This allowed to identify areas of the manipulator's workspace within which no feasible solutions to the above optimization problem exist and hence, no smooth transitions between consecutive trajectory portions are possible. Finally, a motion planning strategy using a superposition scheme allowing to modify ongoing orientational trajectories rather than translational trajectories, under conditions of switching orientation targets, was developed. For real-time applications, the superposition algorithm must be executed quite rapidly since the re-calculation of the trajectory must be performed "on the fly", without interrupting the robotic motion. While the study by Gat-Falik (1990) has indicated the feasibility of applying the superposition strategy to robot manipulators, the search for the optimal values of the switching and movement times has proven to be too time-consuming and not applicable for real-time systems. Recently, a rapid algorithm for searching for the optimal s and T_2 was developed (Rogozin, 1993). This algorithm defined some domain within which the optimal values of these parameters might be found. The domain is reduced as much as possible in order to limit the possible range of values of these parameters that must be checked. The proposed algorithm was implemented on the Adept-one industrial robot and its effectiveness was demonstrated (Rogozin, 1993).

3. MOTOR EXECUTION AND ADAPTATION

To execute any desired motion plan, either in the case of human or robotic arms, appropriate joint torques and actuator forces must be generated. One possible way that this can be done is by first transforming the desired hand or end-effector trajectories into appropriate joint rotations by solving the inverse kinematics problem. Then, the necessary joint torques can be derived by solving the "inverse dynamics" problem. Given the complicated dynamic interactions that exist between the moving limb segments during multijoint movements and the need to distribute the resultant joint torques among the highly redundant sets of muscles that operate about the different joints, it was argued that it is quite unlikely that the biological motor control system explicitly solves these computational problems. Thus, it was suggested that the nervous system must have developed alternative means for motor execution that do not involve explicit joint torque and muscle force computations. One such scheme was proposed in the context of the *equilibrium trajectory* control model (Bizzi *et al.*, 1992). According to this model, the viscoelastic properties of muscles play an important role in allowing the motor system to bypass the need for explicit torque computations. Thus, movements are generated by gradually shifting the limb equilibrium position, along the desired motions, while the equilibrium positions are internally coded by specifying the appropriate neural activities to sets of agonist and antagonist "spring-like" muscles. Alternatively, recent progress in neural network research has led to a renewed interest in the possibility that the motor system does not solve from scratch the inverse kinematics and dynamics problems each time a new movement is about to be generated but that successful solutions to these problems are embedded into the synaptic connections between the elements of the biological cortical and subcortical neural networks that take part in movement generation.

At present there is still a controversy in the motor control literature with respect to the question of how the motor system executes desired motion plans. This issue is closely related to the question of what the overall organization of the motor control system is. One basic notion concerning the latter is that the motor system is hierarchically organized and hence, desired behavioral goals are gradually transformed from more central neural representations that deal mainly with the kinematic aspects of the movements into more downstream representations that are more concerned with motor execution, i.e., with the generation of appropriate joint torques and/or muscle forces (Flash, 1990). According to an alternative point of view, however, movement generation does not involve such step-by-step transformations. This point of view was expressed, for example, by a recent model of human trajectory generation which was again based on optimization theory (Uno *et al.*, 1989). By contrast, however, to the minimum-jerk model, this model assumes that movement selection is based on the optimization of the rate of change of actuator efforts, e.g., expressed as joint torques. Although minimizing jerk and minimizing the rate of change of joint torques appear conceptually similar, there are important differences. First, the latter objective function is based on dynamic variables, therefore the predicted motion depends sensitively on the (assumed) dynamic behavior of the musculoskeletal system. In that sense, this theory is based on the hypothesis that motor computations are executed in parallel, taking both dynamic and kinematic factors into account simultaneously. Secondly, the objective function was formulated in terms of joint torques. This implies that motor computations are based on a joint-space representation of behavior. Uno *et al.* (1989) reported that the performance of the minimum-torque change theory surpassed that of the minimum-jerk model in that it better accounted for several characteristics of human arm trajectories. In particular, the torque-change model predicts that for movements performed in the horizontal plane, all point-to-point hand paths are roughly straight except for movements where the starting point of the arm is at the side of the body and the endpoint is in front of the body. The authors reported experiments confirming these predictions. However, an independent study (Flash, 1990) has shown that the the straightness of the horizontal planar point-to-point movements predicted by Uno *et al.* (1989) is a consequence of unrealistic inertial parameters used to model the human arm. Using more realistic inertial parameters, the trajectories predicted by the minimum-torque change model were found to be unrealistically curved with double-peaked rather than single-peaked speed profiles. The most critical comparison of these two models arises from their fundamental differences. According to kinematically-based optimization models, neural computations specify intended motions independently of movement dynamics or external load conditions. In contrast, dynamically-based optimization models imply that external loads profoundly influence intended motions. Thus, these two alternative models make substantially different predictions about motor behavior following load adaptation.

3.1. *Motor Adaptation Studies*

The question of what is optimized in arm movements was investigated recently from motor learning and adaptation perspectives. Investigating motor adaptation to elastic loads, Uno *et al.* (1989) have concluded that the behavior in the presence of the load is different from the one seen in the unloaded case. Completely different results, however, were obtained by Gurevich (1993) when static elastic loads were unexpectedly introduced during human reaching toward visual targets. Thus, while in the first few trials following load application the movements were found to be misdirected and to miss the final target, following a relatively small number of practice trials (5-7), the loaded movements tended to converge toward the ones seen in the unperturbed case, i.e., to follow straight hand paths with symmetric velocity profiles and thus to converge to the type of trajectories predicted by the minimum-jerk model (Gurevich, 1993). Similar results were observed in another recent study (Shadmehr and Mussa-Ivaldi, 1994) involving the application of velocity-dependent force fields. Thus, it seems that human arm trajectories obey the same kinematic plan independently of the external force conditions thus supporting the idea that the desired behavior is independent of movement dynamics. This conclusion was also supported by the results of a third related study, (Wolpert *et al.*, 1994) in which aiming movements were performed under altered visual feedback conditions involving artificial manipulations of the perceived curvature of the movements. The results from the latter study have also suggested that arm trajectories are planned in extrinsic visual space and are incompatible with assumptions of the minimum-torque change model or of similar models assuming movement generation based on the optimization of dynamics- based cost functions or ones which depend on intrinsic coordinates.

What is it that the CNS learns during skill acquisition and when adapting to new external conditions? To account for the kinematic characteristics of the movements performed in the course of load adaptation Gurevich (1993) has suggested a load adaptation scheme based on the equilibrium trajectory model. According to this scheme, motor adaptation involves the modification of both the arm impedance and the equilibrium trajectory whereby following practice the adapted stiffness and equilibrium trajectory result from the summation of static components needed to overcome the load and "phasic" components responsible for driving the arm along the desired trajectories.

4. NEURAL NETWORK MODELS OF MOTOR LEARNING

In a recent study (Jordan *et al.*, 1994) a somewhat different approach to modeling point-to-point reaching movements than the one used in the minimum-jerk model was taken. In that study, the motor control system was assumed to prefer motions of the hand along straight paths in space. In attempting, however, to account for the temporal aspects of the movements, the characteristic velocity profiles were hypothesized to implicitly arise from the dynamical properties of the neuromuscular hardware, namely the arm and muscle dynamics, including muscle activation–contraction properties, and the dynamics of the neural networks that transform movement commands into muscle activations. That model also suggested that the motor system may possess forward models of the musculoskeletal plant allowing it to deal with a fundamental problem in motor learning theory, namely the "distal teacher" problem (Jordan and Rumelhart, 1992). This problem is concerned with the nature of the corrective feedback that is available to the learner, namely in many motor learning problems corrective information is not provided directly to the learner in terms of motor command errors. Rather, desired behavior is specified in terms of the outcome of movement as assessed by various sensors. One possible way for transforming a distal sensory error (e.g., in the limb spatial position) into signals for correcting the proximal motor commands is by making use of forward models of the kinematics and dynamics of the controlled plant (Jordan and Rumelhart, 1992). The role of the forward models is essentially that of providing a mechanism for transforming distal sensory errors into proximal motor errors. The forward models must themselves be learned by correcting the error between predicted sensory outcomes and actual sensory outcomes. Once those models have been partially trained they can be used to train the controller. The composite system, composed of the controller and the forward models is trained based on the difference between the desired sensation and the actual sensation. While the composite system is being trained the forward models are held fixed and only the controller is altered.

Based on the above ideas, a neural network model was implemented (Jordan *et al.*, 1994) which generated neural inputs to a set of muscles, thereby producing two-joint arm movements. The model consisted of a cascade of neural networks. The first one serving as a *central controller* network and receiving as inputs the initial and the desired final hand positions. It was followed by neural networks serving as forward models of the inverse kinematics and dynamics computations, finally producing hand position outputs which were fed back to the central controller, thus providing recurrent feedback that drove the network forward in time. Various architectures were examined for the central controller. In the final design, the network was composed of one input layer, one output layer and two layers of hidden units with recurrent loops onto themselves and direct connections from the input layer and from both hidden layers to the output layer. The forward model networks were separately trained. The central network was then trained by providing the overall network with inputs only at the first time step, allowing the system to run for the entire duration of the movement. Because the controller and the musculoskeletal system are dynamical systems, errors at the final time step can be caused by poor choices of actions at any of the earlier time steps. To deal with this temporal credit assignment problem, the "backpropagation-in-time" algorithm was used. Errors were passed backward through the forward models and the controller at each time step and the weights in the controller but not in the forward models were changed at the end of this process so as to minimize the chosen performance error.

The network was trained to generate several point-to-point arm trajectories in the horizontal plane and the effects of using several types of performance errors were tested. We found that requiring of the arm to reach the target with prescribed end-point position, velocity and acceleration is inadequate for the purpose of generating human-like trajectories in the case of two-joint planar movements. Thus, an algorithm for training the network based on the deviations of the hand path from a straight line was developed. This algorithm was consistent with our assumption that the training data in learning to reach specifies the desired *paths* of movements, not their *trajectories*. The network was trained to generate a given set of arm trajectories connecting several initial and final targets. At its final design stage and following training it was found that the entire network generates smooth biological-like movements and that the neural commands to the muscles obey the same temporal activation patterns to those experimentally observed in human subjects. Thus, our results have indicated the possibility that the temporal properties of human arm trajectories might be constrained by two factors: the explicit constraint of producing straight-line paths of motion and the implicit smoothing properties of the dynamical systems underlying the movement.

5. CONCLUSION

This article has discussed several biological motion control models dealing with arm trajectory planning and modification, motor execution and motor adaptation. We have also presented several examples of how ideas derived from the study of biological motor control might be applied to deal with motion planning and modification for robot manipulators. Finally, neural network modeling work that has emphasized the significance that internal forward models of arm kinematics and dynamics may play in motor generation and learning was presented. This work may further be extended to deal with other aspects of motor learning and with various robotic and biomedical applications including the development of adaptive controllers for neuroprostheses.

6. REFERENCES

Bizzi, E., N. Hogan, F.A. Mussa-Ivaldi and S. Giszter (1992). Does the neurvous system use equilibrium-point control to guide single and multiple joint movements. *Behavior and Brain Science* **15**, 603-613.

Flash, T. (1990). The organization of human arm trajectory control.. In: *Multiple Muscle Systems. Biomechanics and Movement Organization* (J. M Winters and S. L. Y Woo, Eds.). pp. 283-301. Springer-Verlag. New York.

Flash, T. and E. Henis (1990). Arm trajectory modification during reaching toward visual targets. *J. Cogn. Neurosci.* **3**, 220–230.

Flash, T. and N. Hogan (1985). The coordination of the arm movements: an experimentally confirmed mathematical model. *J. Neurosci.* **7**, 1688–1703.

Flash, T. and R.B. Potts (1988). Discrete trajectory planning. *Int. J. Robotics Res* **7**, 48–57.

Gat-Falik, T. (1990). The superposition strategy for trajectory modification: extensions to robot dynamic tasks. Master's thesis. Dept. of Appl. Math and Computer Sci Weizman Inst. of Science.

Gurevich, I. (1993). Strategies of motor adaptation to external loads during planar two-joint arm movements. PhD thesis. Dept. of Applied Mathematics & Computer Sc., The Weizmann Institute of Science.

Henis, E. and T. Flash (1994). Mechanisms underlying the generation of averaged modified trajectories. *Biol. Cybern., In Press.*

Jordan, M.I. and D.E. Rumelhart (1992). Learning with a distal teacher. *Cognitive Science* **16**, 307–354.

Jordan, M.I., T. Flash and Y. Arnon (1994). Learning arm trajectories from spatial deviations. *J. Cogn. Neurosci.* **6**, 359–376.

Morasso, P. (1981). Spatial control of arm movements. *Exp. Brain Res.* **42**, 223–227.

Rogozin, V. (1993). Robot control for dynamic tasks. Master's thesis. Dept. of Appl. Math and Computer Sci Weizman Inst. of Science.

Shadmehr, R. and F. A Mussa-Ivaldi (1994). Adaptive representation of dynamics during learning of a motor task. *J. Neurosci.* **5**, 3208–3224.

Uno, Y., M. Kawato and R. Suzuki (1989). Formation and control of optimal trajectory in human multijoint arm movement: Minimum torque-change model. *Biol. Cybern.* **61**, 89–101.

Wolpert, D.M., Z. Ghahramani and M.I. Jordan (1994). Are arm trajectories planned in kinematic or dynamic coordinates? an adaptation study. *Exp. Brain Res., In Press.*

LEARNING TO OPTIMIZE PERFORMANCE: LESSONS FROM A NEURAL CONTROL SYSTEM

Chi-Sang Poon

Harvard-MIT Division of Health Sciences and Technology
Massachusetts Institute of Technology
Cambridge, MA

Abstract: Cognitive processes in human-machine interaction are mostly executed in the cerebral cortex (and other parts of the higher brain) which is designed for multi-tasking operations but is generally not optimized for any specific task. Similar cognitive and control operations also occur in the brain stem which is composed of many task specific neural modules for the homeostatic regulation of a host of physiological functions using peripheral processes as the "machine". In the respiratory control system, the brain stem controller demonstrates intelligent control capability and achieves optimal performance under a wide variety of environmental conditions. Such intelligent control behavior can be realized by reinforcement learning in an associative Hebbian synapse in the brain stem. Similar reinforcement learning modules in the higher brain can be fashioned by repeated training to effect robust optimal operation in a human-machine system.

Keywords: Neural control; learning systems; decision making; optimal control; performance indices; physiological models

1. INTRODUCTION

One of the main objectives in designing a human-machine system is to optimize the operation of the overall system. The performance of any human-machine system is dependent in part on the coordination between the machine and the human operator. The behavior of the machine is often preset and is dictated by the specific applications concerned. It is therefore incumbent upon the operator to learn and adapt to the human-machine interface in order to achieve the desired performance. Because of their innate adaptability, human operators can usually operate under a wide variety of environmental constraints. Even so, human adaptation is limited by many physical and cognitive factors. In order to optimize the design of a human-machine system it is necessary to understand the mechanism of learning and adaptation processes in the brain and then design a human-machine interface that conforms to such learning behavior.

In this paper we examine the learning and optimization behavior of a specific neural control system - the respiratory control system - as an example of how the brain might learn to optimize the performance of a human-machine system.

2. COGNITIVE AND CONTROL PROCESSES IN THE BRAIN

2.1 Neural computations in the higher brain

Human decision making is a complex biological process that involves numerous intricate and concerted mental computations. It is not clear how such mental computations come about in the brain to reach an optimal decision for a particular task. Mental decision may represent an emergent behavior that requires the collective action of many unit processes.

In addition to the human-machine task at hand the brain is responsible for a host of other physiological functions that take place simultaneously. The cerebral cortex of the brain - which is the central "controller" of human volitional behavior - may be thought of as a general-purpose computer that subserves many different life-sustaining programs running simultaneously in parallel (Fig. 1). The advantages of such a general parallel computation architecture are obvious: it renders the brain immense computing power with a great deal of flexibility, adaptability and efficiency. This is why human operators can handle multiple tasks simultaneously and can learn a difficult task or adapt to a novel work environment readily.

Such flexible, multi-tasking and large-scale computation capabilities of the cerebral cortex are its key assets. However, in some circumstances such unique qualities of the cerebral cortex may also become its disadvantages. Thus, in performing multiple tasks in parallel the brain is subject to distraction or cross talk. Although a generalized computation architecture may be desirable because of its adaptability to many different tasks, it may not be optimized for any particular task (i.e., the neural analog of "Jack of all trades"). An ideal human-machine interface should therefore provide a work environment that would take the greatest advantage of the operator's mental and physical capacities.

2.2 Neural computations in the lower brain

As a first step toward designing an effective human-machine system, it is necessary to understand how the brain makes optimal decisions when subjected to varying environmental constraints. Unfortunately, the complexity of the cerebral cortex and its ancillary brain structures makes it extremely difficult to elucidate the mental processes that lead to optimal decision making in the higher brain. Instead, one may gain a great deal of insight by studying a simpler neural system that exemplifies such decision and control processes.

Aside from volitional activities, the brain is also engaged in a wide variety of decision and control functions at the subconscious level. These include the automatic regulation of many vital physiological functions such as circulation, respiration, body temperature, etc. These physiological functions are largely involuntary and are regulated by specialized neuronal groups in the brain stem and other regions of the lower brain (Fig. 1). Unlike volitional decision

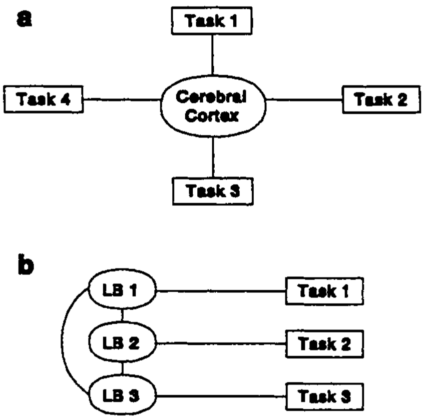


Fig. 1 The brain as a controller. (a) The cerebral cortex as a general purpose, multi-tasking parallel computer. (b) Brain stem modules (LB) as dedicated, task-specific microprocessors connected in parallel.

making by the cerebral cortex, such regulatory processes operate largely in an auto-pilot mode at the subconscious level with little voluntary intervention necessary. And yet, just like the higher brain, these subconscious neural structures display remarkable adaptability and intelligence in decision making and control under varying physiological environments. Indeed, it is such adroitness of physiological control systems that allows the body to maintain homeostasis, or constancy of internal environment, under a wide range of internal and external disturbances to physiological functions.

As an example of such intelligent decision and control functions in the brain, we will examine the neural computations involved in one such neural control systems - the respiratory control system.

3. THE RESPIRATORY SYSTEM AS A HUMAN-MACHINE SYSTEM

3.1 Mechanics and kinetics of the respiratory system

A prime function of the respiratory system is to sustain body metabolism by exchanging metabolic gases, O_2 and CO_2. This is carried out by the act of breathing - the rhythmic contraction and relaxation of the respiratory muscles - which moves air in and out of the lung. The respiratory process therefore takes place in two steps. First, the "operator" must act on a "mechanical plant" (the lung and chestwall) by exertion of the respiratory muscles to expand the lung. This is given by the following equation of motion:

$$R_{rs} \cdot \frac{dV}{dt} = -E_{rs} \cdot V + P_{aw}(t) - P_{mus}(t) \quad (1)$$

where V is lung volume expansion; R_{rs} and E_{rs} are the total resistance and elastance of the lung and chestwall; P_{mus} is the pressure generated by the respiratory muscles; and P_{aw} is any external pressure to facilitate respiratory air movement (for example, by means of a mechanical respirator).

Second, the resulting air movement (dV/dt) due to lung expansion and relaxation then ventilates the "chemical plant" (the lung) to effect pulmonary gas exchange. The effective pulmonary ventilation, $\dot{V}_E$, defined as the mean airflow in and out of the lung, determines the CO_2 tension in the arterial blood (P_aCO_2) according to the following kinetic equation:

$$V_L \cdot \frac{d}{dt} P_aCO_2 = \dot{V}_E(P_ICO_2 - P_aCO_2) + 863\,\dot{V}CO_2 \qquad (2)$$

where V_L is the CO_2 store in lung tissues; P_ICO_2 is the partial pressure of CO_2 in the inspired air; and $\dot{V}CO_2$ is the whole-body metabolic production of CO_2.

The combination of the chemical-mechanical plant of the lung-chestwall system constitutes the "machine" that one operates in order to breathe.

Ordinarily, breathing is an involuntary act controlled by specialized neurons in the brain stem which form the respiratory controller. Despite its involuntary nature the respiratory control system, as a human-machine system, works remarkably well. Indeed, nothing seems easier than breathing itself: one seldom has to worry about it and yet it never fails as long as one is alive. Such a human-machine interface remains virtually seamless even when subjected to large environmental and physiological perturbations.

The respiratory system is a good example of a human-machine system that is designed by nature for optimal performance. To better characterize the behavior of this system we will first define the performance measure that is implicit in the physiological task of breathing.

3.2 *Performance measure*

Everyday experience suggests that one usually does not breathe any harder than what is demanded by the body. Numerous laboratory tests have shown that during muscular exercise where body metabolism is increased, respiratory output is closely matched to the metabolic demand of the body. In the steady state, the increase in $\dot{V}_E$ is always proportional to the increase in $\dot{V}CO_2$. Consequently, homeostasis of P_aCO_2 is closely maintained throughout exercise, as suggested by Eq. 2.

However, such homeostasis is abolished when, instead of breathing room air where $P_ICO_2 \approx 0$, one rebreathes CO_2 (e.g. in a crowded and underventilated room) thereby clogging the lung and airways. In this event the efficiency of gas exchange in the lung is so severely compromised that it is impossible to restore the normal P_aCO_2 no matter how hard one tries to breathe. This can be seen from Eq. 2 where the condition $P_aCO_2 \geq P_ICO_2$ holds for all $\dot{V}_E$. Experimental observations suggested that in this case the respiratory controller behaves quite differently than during muscular exercise: it also increases respiratory effort in response to increased CO_2 load, but the resulting increase in $\dot{V}_E$ is not enough to totally eliminate the CO_2 load from rebreathing. Consequently, P_aCO_2 tends to rise thus disrupting the homeostasis found during muscular exercise.

Thus we have an interesting experimental paradigm where the controller seems to differentiate between the sources of CO_2 load to the lung and respond totally differently in different circumstances. In the case of metabolic CO_2 produced by muscular exercise the respiratory controller maintains absolute stability of P_aCO_2 by increasing respiratory effort, whereas for any rebreathed CO_2 the controller allows P_aCO_2 to rise without over-exerting respiratory effort (Fig. 2). Such a discrepancy in response behavior with different types of CO_2 load has remained an enigma in physiology and has been a subject of extensive research for over a century (Poon, 1995).

One possible explanation of such discrepancy in response behavior is that the object of breathing is not merely to replenish oxygen and eliminate carbon dioxide. Otherwise, one would have to keep breathing very hard to ventilate the lung as much as possible. In so doing one would be incurring a lot of energy expenditure since breathing is a repetitive motor act that requires appreciable work done continually. Therefore, it is reasonable to assume that a secondary object of the respiratory system is to minimize the work rate of breathing as much as possible.

Minimization of work rate is often an important object in any human-machine system since

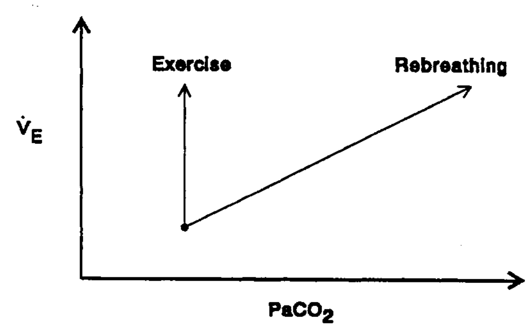


Fig. 2 Differential responses of the respiratory controller to CO_2 load from exercise and rebreathing.

mechanical work done is a major physical constraint of the operator and may become a limiting factor in a human-machine system. Thus the operator is always confronted with two conflicting objectives: to accomplish the primary task by performing the requisite work on the one hand, and on the other hand to minimize the work done as much as possible in order to conserve energy.

For the respiratory system the primary task is to maintain chemical homeostasis by the motor act of breathing, and the secondary goal is to minimize the work needed to accomplish this task. These are the chemical and mechanical objectives of the respiratory system, denoted by J_c and J_m, respectively (Poon, 1987; Poon et al., 1992):

$$J_c = \alpha^2 (P_a CO_2 - \beta)^2 \qquad (3)$$

$$J_m = \ln \dot{W} \qquad (4)$$

where α, β are sensitivity and threshold parameters and $\dot{W}$ is the work rate of breathing. Equation 3 measures the chemical penalty due to deviation of blood chemistry from the nominal (homeostatic) level. Equation 4 measures the mechanical penalty associative with the work of breathing which is expressed as a logarithmic function in agreement with the response characteristic of sensory perception.

The overall objective function of the respiratory system is given by the sum of the chemical and mechanical penalties:

$$J = J_c + J_m \qquad (5)$$

The goal of the respiratory system is to minimize this compound objective function subject to the constraints imposed by the chemical and mechanical plants of the lung-chestwall system (Eqs. 1 and 2). Note that J_c and J_m are conflicting objectives since an increased respiratory effort would tend to *decrease* J_c and *increase* J_m.

The compound objective function J represents a balance between opposing physiological factors that are essential to survival. An excessive increase in J_c is deleterious to the body because it disrupts normal blood chemistry, but a large increase in J_m is also undesirable because it causes too much stress to the respiratory muscles and interferes with other activities such as speech and coughing.

4. OPTIMIZATION BEHAVIOR OF THE RESPIRATORY SYSTEM

Given the above operational constraints, the respiratory controller must try and optimize its performance by carefully balancing the chemical and mechanical penalties associated with breathing. This is a difficult computational task because it amounts to solving a nonlinear, multi-input multi-output dynamic optimization problem (Eqs. 1 - 5). Moreover, this must be carried out continuously by on-line neural computation since the body is subject to continuous environmental and internal disturbances.

4.1 Optimal performance of the respiratory system under disturbances to chemical plant

With the performance measure defined in Eqs. 3 - 5 one can evaluate the performance of the respiratory system by comparing the optimal responses predicted by the model equations to those observed experimentally. Experimental data showed that the respiratory controller performs remarkably well in balancing the chemical and mechanical penalties of breathing.

The optimal response is obtained by setting $\partial J/\partial \dot{V}_E = 0$ subject to the constraint given by Eq. 2 and the relationship $\dot{W} \propto \dot{V}_E^2$, yielding (Poon, 1987; Poon et al. 1992):

$$\dot{V}_E = \alpha^2 (P_a CO_2 - \beta) \dot{V}CO_2 \qquad (6)$$

Thus optimal performance of the respiratory system, as defined by the performance measure in Eq. 5, calls for a proportional relationship between $\dot{V}_E$ and $\dot{V}CO_2$ at constant $P_a CO_2$ during muscular exercise and a linear relationship between $\dot{V}_E$ and $P_a CO_2$ during CO_2 rebreathing at constant $\dot{V}CO_2$. These response behaviors are similar to those observed in the respiratory system (Fig. 2).

This finding suggests that the homeostasis of blood gases and pH in the body during muscular exercise is optimal for the respiratory system. More important, the model suggests that disruption of homeostasis during CO_2 rebreathing may also represent an optimal response in the sense that the controller exchanges the chemical penalty for a reward in conservation of respiratory work. In both cases, the overall objective function J is minimized under the corresponding operational constraints.

Such an ability of the respiratory controller to optimize its performance under varying operational constraints is characteristic of many physiological processes that are regulated by brain stem neurons. As another example, in healthy individuals the body temperature is well regulated at 37° C, but under severe temperatures or in a fever where body function is under attack, the homeostasis of body temperature is abolished in favor of conservation of energy or facilitation of healing from disease.

4.2 Optimal performance of the respiratory system under disturbances to mechanical plant

Respiration may also be interrupted if the mechanical plant is disturbed. This occurs, for example, with a common cold or in lung disease where the airways are clogged and the lung becomes less compliant, making it harder to breathe. Common experience shows that the body automatically defends against the increased mechanical challenge by augmenting respiratory effort. Thus, pulmonary ventilation is restored and chemical homeostasis is maintained except in severe disease states (e.g. asthma).

A greater mechanical load increases the energetic penalty J_m for breathing. Thus, it becomes even more important to conserve energy by minimizing the work of breathing. However, this must be balanced against possible increases in chemical penalty. Rather than reduce respiratory effort and $\dot{V}_E$, which would jeopardize pulmonary gas exchange and incur a chemical penalty, the respiratory controller is found to optimize the breathing pattern by fine tuning the inspiratory and expiratory waveshapes in compensation for the mechanical load. Numerous studies have shown that the waveshapes and patterns of respiratory motor output are consistent with the optimal waveshapes found by the minimization of $\dot{W}$ under varying mechanical loads (Poon et al., 1992).

4.3 Optimal patient-respirator interaction

Normally, breathing occurs when the respiratory muscles operate on the lung-chestwall system which may be considered as a passive "machine" defined by Eq. 1 with $P_{aw} = 0$. In respiratory failure, the spontaneous driving pressure P_{mus} is not strong enough to generate adequate respiratory airflow, resulting in hypoventilation. This clinical condition may be caused by neuromuscular disease in which P_{mus} is weak, or pulmonary disease in which R_{rs} and/or E_{rs} become very large. In both cases, a mechanical respirator is necessary to maintain pulmonary ventilation by providing an external driving pressure, P_{aw}. The respirator is an active machine which interacts continuously with the patient to produce pulmonary ventilation (Fig. 3).

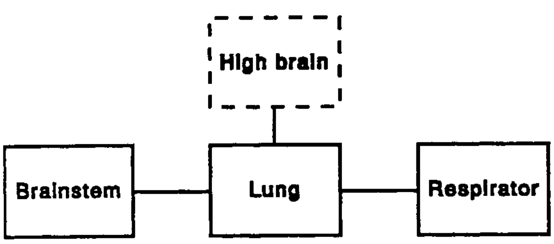


Fig. 3 Patients adapt to a respirator more readily when cortical influence is suppressed by sedation.

Conventional respirators are generally designed to cycle automatically according to a preset rhythm as prescribed by the physician. However, the preset mechanical rhythm may not always coincide with the patient's spontaneous respiratory rhythm which may fluctuate continuously depending on the physiological and psychological states. In awake patients, the presence of a respirator often provokes anxiety and discomfort. Consequently, some patients may "fight" the respirator creating a problem in mechanical ventilation. The clinical treatment for such patients often calls for sedation or muscle relaxants to eliminate erratic respiratory efforts.

The fact that sedated or anesthetized patients are generally more receptive to mechanical ventilation than awake patients suggests that erratic respiratory responses to respirator treatment is mainly a psychological reaction originating from the higher brain. In contrast, the respiratory controller in the brain stem appears to be highly adept at interacting with the respirator in maintaining pulmonary ventilation (Fig. 3).

The anxiety associated with mechanical ventilation may be alleviated by synchronizing the respirator to the patient's spontaneous respiratory effort. In this event patients are often found to take advantage of the mechanical assistance from the respirator by lowering his/her respiratory effort. As a result, pulmonary ventilation is maintained at the normal level but with substantially reduced work of breathing (Poon et al., 1987). This behavior is again in agreement with the model prediction (Eqs. 1, 5) in the presence of mechanical assistance to respiration, indicating optimal performance of the patient-respirator system.

Furthermore, it has been shown that when the assistance from the respirator becomes excessive, the respiratory controller may fight the hyperventilatory effect in a very intelligent way. Rather than directly oppose the respirator pressure, which may cause excessive back pressure in the lung resulting in barotrauma, the controller is found to generate a respiratory rhythm that is 180° out of phase with respect to the respirator rhythm. Such a strategy proves to be very effective in countering the hyperventilatory effect of the respirator without causing excessive mechanical damages to the system (Poon and Kolandaivelu, 1994).

In summary, these experimental findings suggest that the respiratory system is an intelligent control system which maintains optimal performance under a wide variety of human-machine interactions. It is of interest to understand how such optimal performance is achieved by the respiratory controller. Such information may provide insight into the design of human-machine systems with similar operational constraints.

5. REINFORCEMENT LEARNING IN RESPIRATORY CONTROL

5.1 Learning and optimization in the brain stem

Optimization of the performance measure (Eq. 5) amounts to solving two-point boundary value problems continuously on-line in the presence of time-varying disturbances. How does the controller deal with such difficult computational problems?

The changes in response behavior from one physiological state to another (Fig. 2) suggest that adaptive neural mechanisms may be involved in respiratory optimization. One possible mechanism of adaptive neural control is reinforcement learning in a Hebbian synapse with persistent excitation (Poon, 1994). In particular, a correlational form of Hebbian synaptic plasticity has been proposed to be a possible neuronal substrate for respiratory optimization with an adaptation rule of the form (Poon, 1993):

$$\delta G = -k[\delta P_aCO_2 \cdot \delta A + G\cdot(\delta \ln A/A_0)^2] \quad (7)$$

where G is the synaptic weight which determines the throughput gain of the controller; A is the activity of the output neuron and is directly proportional to $\dot{V}_E$; and k is a constant.

The synaptic adaptation rule in Eq. 7 is a neuronal realization of reinforcement learning because the synapse is strengthened/weakened by any negative/positive correlations in the input-output relationship ($\delta P_aCO_2.\delta A$). Thus, the controller gain is increased if any increase in respiratory activity is rewarded by a corresponding lowering in P_aCO_2; i.e., an increase in mechanical work is compensated for by a greater decrease in chemical penalty (Eq. 5). This neuronal model has been shown to conform with the optimization of the chemical and mechanical objectives of respiration (Poon, 1993).

5.2 Learning and optimization in the higher brain

If reinforcement learning is indeed how the respiratory system optimizes its performance, then the brain stem seems to be remarkably adept at learning control. To examine whether the higher brain is capable of such optimization tasks, an experiment was performed in which human subjects were asked to interact with a specially designed machine which was programmed to simulate the dynamics of the respiratory system (Eqs. 1 - 2) so as to minimize an objective function (Eq. 5) which is continuously calculated and displayed on an oscilloscope screen. Thus the machine simulates the lung-chestwall system while the subject's visuomotor system becomes a "proxy" of the respiratory controller. After some trial and error, subjects were found to be able to approximate the output solutions corresponding to different types of CO_2 inputs (Fig. 2), but the responses were less than optimal compared to those achieved by the brain stem controller under similar simulation conditions (Poon, 1991).

CONCLUSIONS

Several useful lessons about human-machine interaction can be learnt from the respiratory control system. First, it is possible for a human operator to optimally adapt to a machine through on-the-job reinforcement learning. Second, in order to do so the reinforcement signal must strongly and directly influence mental decisions of the operator without needing much decoding. Third, persistent excitation of the system is necessary for reinforcement learning in time-varying environments. Finally, optimal performance may be attained only if the operator is well trained and dedicated to the specific human-machine task.

ACKNOWLEDGEMENT

This work was supported in part by NSF grant BES-9216419, ONR contract N00014-95-0414, and NIH grants HL45261, HL50614.

REFERENCES

Poon, C.-S. (1987). Ventilatory control in hypercapnia and exercise: optimization hypothesis. *J. Appl. Physiol.* **62**:2447-2459.

Poon, C.-S. (1991). Optimization behavior of brainstem respiratory neurons: a cerebral neural network model. *Biol. Cybern.* **66**:9-17.

Poon, C.-S. (1993). Adaptive neural network that subserves optimal homeostatic control of breathing. *Annals Biomed. Engr.* **21**:501-508.

Poon, C.-S. (1994). Hebbian synaptic plasticity: a neural mechanism of supervised learning and adaptive control. In: *Proc. IFAC Symposium on Modeling and Control in Biomedical Systems* (B.W. Patterson (Ed.)), pp. 497-500.

Poon, C.-S. (1995). Respiratory models and control. In: *Biomedical Engineering Handbook, Sect. XV. Physiological Modeling, Simulation and Control* (H.J. Chizeck (Ed.)), pp. 2394-2411. CRC Press, Boca Raton, Fl.

Poon, C.-S., S.L. Lin and O.B. Knudson (1992). Optimization character of inspiratory neural drive. *J. Appl. Physiol.* **72**:2005-2017.

Poon, C.-S. and K. Kolandaivelu (1994). Phase reversal of respiratory rhythm during assisted mechanical ventilation. *Oxford Conf. on Modelling and Control of Ventilation*. London, GB.

PERCEPTION OF COHERENCE OF VISUAL AND VESTIBULAR VELOCITY DURING ROTATIONAL MOTION

Han F.A.M. van der Steen and Ruud J.A.W. Hosman

Delft University of Technology, Faculty of Aerospace Engineering
P.O. Box 5058, 2600 GB Delft, The Netherlands
e-mail: Han.Steen@LR.TUDelft.NL

Abstract. Psychophysical experiments are described that concern the ability of subjects to indicate the velocity coherence of vestibular and visual stimuli during rotations. The coherence zone is determined by the thresholds at which the suggested outside visual world is no longer perceived as stationary. The results show high gains for roll and swing (1 to 6) and gains about one (0.5 to 2) for yaw motions. The use of perceived coherence zones in moving-base vehicle simulation is discussed.

Key Words. Human perception, Visual motion, Thresholds, Many-degrees-of-freedom systems, Simulators

1. INTRODUCTION

While we walk or drive a vehicle, some self-motion sensitive receptors are continuously receiving signals from the environment, while others are registering the state of our internal body. Our nervous system processes these signals and provides information on how we move. With this information, our body generates reflexes that serve, for instance, to maintain balance or to stabilise the image in our eyes. At the conscious level, the sensory information allows us to perceive our present state of motion, and to determine what action should be performed next.

The primary sensory systems that provide information on self-motion are the visual and the vestibular system (Howard, 1982). The visual system provides information about the momentary position and velocity of the body in the environment. The vestibular system is considered to be the most important inertial sensory system for self-motion. Its receptors are sensitive to specific forces and rotational accelerations. In the frequency range of about 0.1 to 10 rad/s, however, rotations induce a vestibular output that is correlated to velocity (Hosman and Van der Vaart, 1978).

In normal conditions, visual and vestibular self-motion information usually is coherent. A visual or vestibular motion stimulus, however, can only be perceived with finite accuracy. Therefore, some mismatch between the stimuli may exist, without affecting the perception of coherence of the observer. For example, motion direction, motion magnitude and time difference of concurrent visual and vestibular motion each may have non-coherence to some extent, while this is not perceived by the observer. The coherence *zone* determines the range of visual and vestibular stimuli that are perceived as being coherent.

This margin of freedom in visual and vestibular motion generation, while being perceived as coherent, can be utilised for the motion control of realistic vehicle simulators. The visual display of a vehicle simulator presents a virtual outside world. The human visual system, however, is apt to take the screen's image as the real visual world. Once this occurs, a situation of non-coherence with the vestibular sense can exist. If the driver becomes aware of such a non-coherence, a so called 'visual-vestibular conflict' or 'false cue' exists. Properly, the problem resides in the interpretation of the artificial visual world as being real.

In this paper, experiments are described which determine the perceived coherence of a briefly presented sinusoidal visual motion with the vestibularly sensed velocity of a sinusoidal inertial motion. In contrast to previously reported experiments on linear motions (Hosman and Van der Steen, 1993), the current work considers rotational motions. The results of these experiments can be used for a new concept on vehicle simulation, using perceived coherence zones.

2. METHODS

In order to determine the coherence zone of velocity magnitude, the physical coherent situation of the inertial and visual stimulation was deliberately mismatched. The motion of the visual stimulus could be manipulated more easily than that of the inertial stimulus. Therefore, an inertial stimulus of equal form and magnitude was induced in each session. The perceived coherence of different visual velocities with the inertial motion was tested.

2.1. *Experimental set-up*

The subject was seated in the cockpit of a simulator base (Figure 1). In the cockpit, a small fixation light (LED) was placed in front of the subject and two monitors were placed in the peripheral visual field (Figure 2). Extraneous light within the cockpit was minimised. The laboratory was darkened and the exterior of the cockpit was covered with heavy cloth. The inside of the cockpit was painted flat black to prevent reflections from the LED and monitors. To mask auditory cues from the moving base, the subject wore a head-set, through which music was played during the sessions.

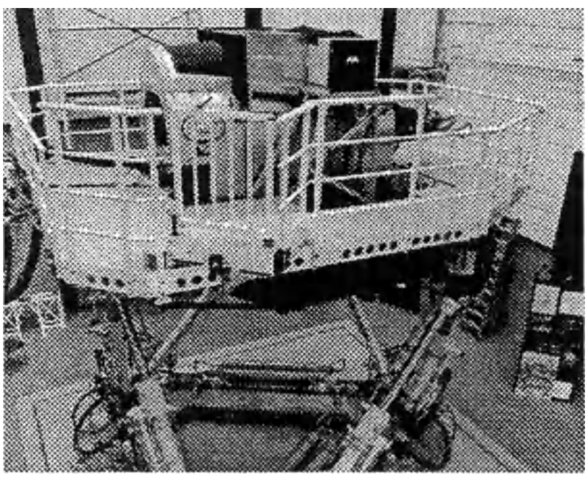


Fig. 1. The NLR 6-DOF research moving base.

2.2. *Motion stimuli*

Three types of motion were tested: roll, swing, and yaw. The rotation of each motion was sinusoidal with fixed frequency ω and amplitude ϕ_i:

$$\phi(t) = \phi_i \sin(\omega t) \qquad (1)$$

In the roll motion, the subject was rotated about the x-axis (Figure 2). In the swing motion, the subject was oscillated as a simple pendulum. The axis of rotation was parallel to the x-axis with a distance z_0 above the centre of the head, equal to

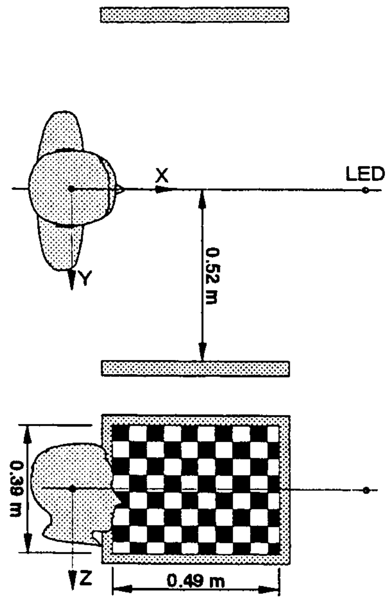


Fig. 2. The visual stimuli configuration for the roll and swing experiment. For the yaw experiment, the monitors were slanted.

the length of the virtual pendulum bob:

$$z_0 = \frac{g}{\omega^2} \qquad (2)$$

where g is the acceleration of gravity. The rotation component of this motion was the same as for roll. The shear specific force on the body that occurred in roll, however, was compensated by side-to-side sway motion in the y-direction and heave motion in the z-direction. In the yaw motion, the subject was rotated about the z-axis.

The peripheral visual stimulus consisted of checkerboard patterns. These patterns were displayed on monitors and represented the rotation of the outside visual world. This stimulus also had a sinusoidal motion profile and had the same frequency and rotation axis as the inertial stimulus. Unlike the inertial stimulus, however, the amplitude was altered after each period. Near the maximum velocity of the inertial motion, the visual motion was presented for 0.40 s. By using such a short presentation time, a visually induced motion sensation (optokinese) was avoided, while retaining the possibility to sense the velocity which is necessary for the determination of coherence. The perceived self-motion was therefore induced only by the vestibular sense.

2.3. *Experimental sessions*

Each session was aimed at the determination of the visual velocities which the subject perceived to be coherent with the inertial velocity, sensed by the vestibular system. These velocities were determined for several frequencies and amplitudes of the inertial motion.

The simulator base was set in sinusoidal motion. The subject was instructed to fixate on the LED to prevent visual motion induced by eye movement. After two oscillations, the session started. The visual stimulus was presented once in each period. The moment of presentation was previously chosen to be at the time when the subject moved either to the right or to the left. The task of the subject was to indicate whether or not the checkerboard patterns were moving exactly opposite to the inertial motion; as it should be for a stationary world. The subject had three buttons to give an answer. One button for perceived visual motion too slow; one for motion too fast; and one for perceived coherence.

2.4. *Coherence thresholds*

The borders of the too-slow and too-fast perceived regions are called the *slow* and *fast* thresholds, respectively. In each session, one of the thresholds was searched for. This was done using a staircase method where the velocity amplitude of the visual stimulus was altered after each period. The step size and search direction of the staircase method depended on the previous answer. At the start, the step size was large; near the threshold the steps were small. The step size was divided by two when search direction was changed. The threshold was found when the step size had become smaller than a value, previously set by the experiment leader. The session also stopped when the number of periods exceeded 30, indicating that the staircase algorithm did not converge.

2.5. *Stimulus range*

A frequency range of 1.0 through 2.0 rad/s was chosen for roll and swing. For yaw, a frequency of 1.2 through 2.0 rad/s was chosen. The base amplitudes varied from 0.032 through 0.191 radians for roll; 0.032 through 0.144 radians for swing; and 0.032 through 0.321 radians for yaw. Consequently, the inertial motions were well above vestibular motion perception threshold (Hosman and Van der Vaart, 1978).

2.6. *Subjects*

Five subjects participated in both the roll and swing experiments. Five other subjects participated in the yaw experiment. The age of the subjects was between 21 and 28.

3. THE CONCEPT OF PMC

A threshold gain can be defined as the ratio of visual and inertial velocity at the visual presentation moment:

$$G = \frac{v_{visual}}{v_{inertial}} \tag{3}$$

From the gains at the slow and fast threshold, the point of mean coherence, PMC, is defined as the mean value of these gains:

$$PMC = \frac{1}{2}(G_{slow} + G_{fast}) \tag{4}$$

The PMC indicates the relative weight of the vestibular and the visual perceived motion signal. If the PMC is only velocity dependent, then:

$$PMC = k\omega\phi_i \tag{5}$$

using the first time derivative of ϕ_i from Eq. 1. Then, at PMCs of equal magnitude:

$$\begin{aligned} \log(\phi_i) + \log(\omega) &= \log(PMC) - \log(k) \\ &= Constant \end{aligned} \tag{6}$$

Writing $\log(\phi_i)$ as a function of $\log(\omega)$, the slope of the iso-PMC contours that satisfy Eq. 6 is -1. If the PMCs are only velocity dependent, results from stimulus combinations with equal velocity may be averaged.

4. RESULTS

The conditions of either right or left presentation of the visual stimulus did not lead to different gains. These thresholds are therefore averaged. Contour plots are composed by interpolating PMCs, calculated by a near-distance algorithm. The measured PMCs are the mean of the results of all subjects. These plots illustrate velocity dependency of the PMC. In threshold gain vs. velocity figures, the results of inertial motions with equal velocity are averaged. In the figures showing the thresholds of each subject, the standard error is drawn. In the figures showing the mean results of the subjects, the group standard deviation is shown.

4.1. *Roll*

Figure 3 shows the group PMCs for roll in a contour plot. Although the slope of the iso-PMC lines is not -1 everywhere in the stimulus domain, averaging PMCs that are measured at equal inertial velocities seems to be allowed. At the highest velocity, there is a strong curvature, but the PMC value does not change much. The thresholds of each subject are depicted in Figure 4. From this figure, it is clear that the gains are far higher than one, the physical coherent situation, even for the slow threshold. Notice the slightly higher gains of subject 1 (S1). The

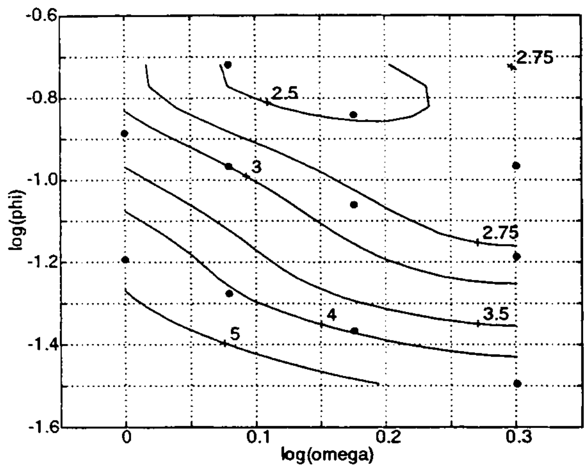


Fig. 3. Interpolated iso-PMC contours for roll. Dots indicate the measured PMCs; numbers indicate the contour value.

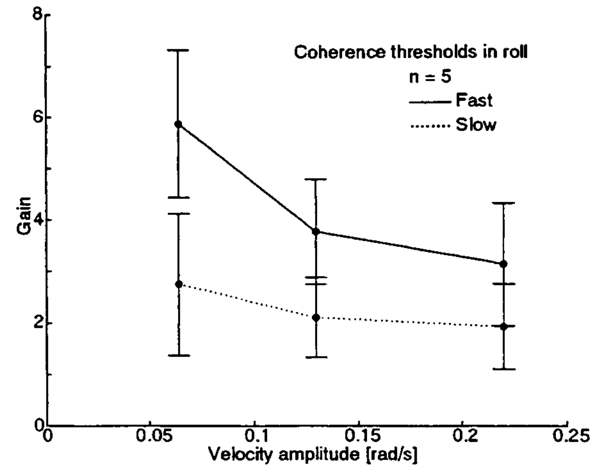


Fig. 5. Group result from the roll experiment. The error bar indicates twice the group standard deviation.

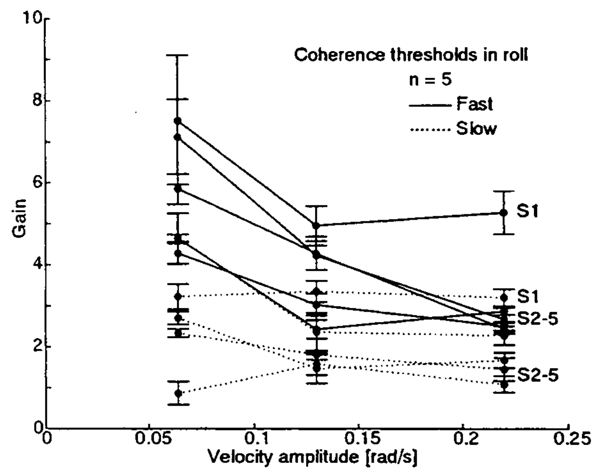


Fig. 4. Results from the roll experiment. The error bar indicates twice the standard error of each threshold.

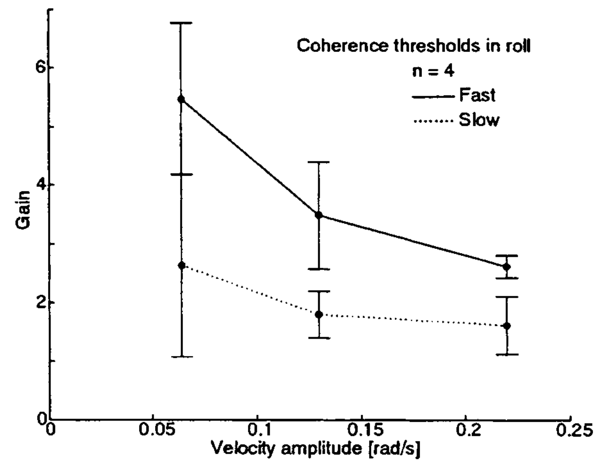


Fig. 6. Group result from the roll experiment, omitting S1. The error bar indicates twice the group standard deviation.

group thresholds are depicted in Figure 5. This figure shows that the gain of the fast threshold decreases from 6 to 3; that of the slow threshold from about 3 to 2. The standard deviation is rather high, mainly due to the gains of S1. For later comparison with the swing experiment, the group result omitting S1 is shown in Figure 6.

4.2. *Swing*

The thresholds of each subject in the swing condition are depicted in Figure 7. Subject 1 (S1) has much higher gains than the rest of the group. From this figure, it can be seen that the gain of the fast threshold is higher than one. Apart from subject 1, the gain of the slow threshold is about one. Figure 8 shows the group PMCs in a contour plot. Whether averaging PMCs at equal inertial velocities is appropriate, is not clear. In the mean group results, subject 1 is omitted, which leads to a low standard deviation (Figure 9). In this figure it can be seen that the gain of the fast threshold decreases from about 4 to 2.5; that of the slow threshold is about 1 for each velocity tested.

4.3. *Yaw*

Figure 10 shows the group PMCs in a contour plot. Note that the axes scale is not the same as that in the previous contour figures. It is clear that PMCs at equal inertial velocities may be averaged. The thresholds of each subject are depicted in Figure 11. Apart from subject 2 (S2), the gains of the fast thresholds decrease to a value below one. The group mean thresholds are depicted in Figure 12. The gain of the fast threshold decreases from 2 to 1 after which a slight increase can be seen, mainly due to subject 2. These higher gains of subject 2 also result in a large group standard deviation at high velocities. The gain of the slow threshold decreases from 1 to 0.5. It is clear that the threshold magnitudes are

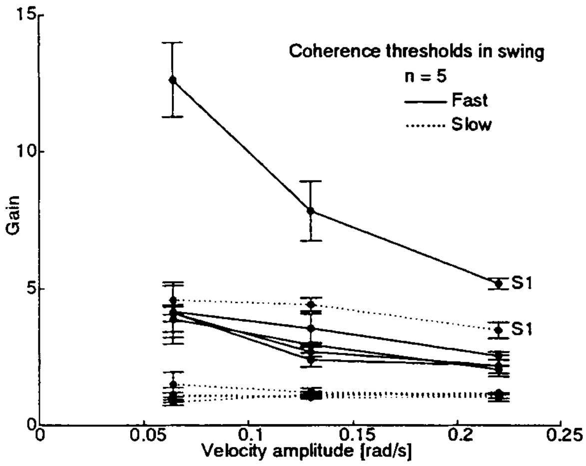


Fig. 7. Results from the swing experiment. The error bar indicates twice the standard error of each threshold.

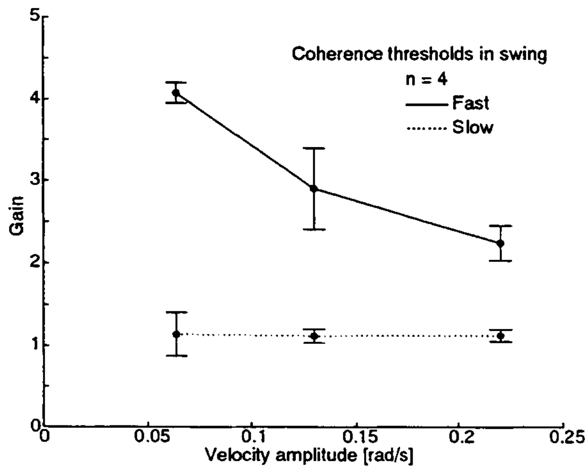


Fig. 9. Group result from the swing experiment, omitting S1. The error bar indicates twice the group standard deviation.

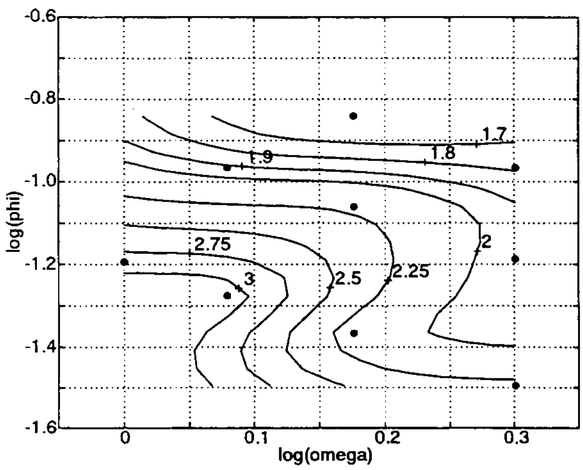


Fig. 8. Interpolated iso-PMC contours for swing, omitting S1. Dots indicate the measured PMCs; numbers indicate the contour value.

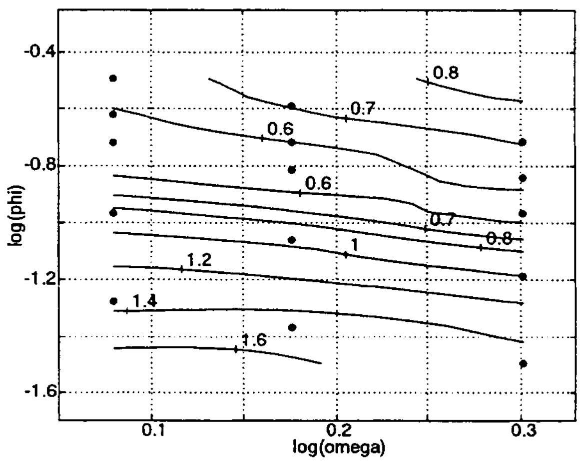


Fig. 10. Interpolated iso-PMC contours for yaw. Dots indicate the measured PMCs; numbers indicate the contour value.

lower than those in the roll and swing conditions, for all velocities tested.

5. DISCUSSION

5.1. *Result discussion*

Velocity dependency of the PMCs is demonstrated clearly for yaw motions, in the stimulus domain tested. For roll and swing, it is likely that velocity is the dependent quantity, although the results are too scattered to come to a definite conclusion.

Although the short presentation time introduces uncertainty in visual velocity estimation, it does not seem to affect the size of the coherence zones much. The zones are much smaller in yaw than they are in roll and swing.

Most threshold gains are well above one, especially for roll and swing. Such high gains are not found in linear motions (Hosman and Van der Steen, 1993; Wertheim and Mesland, 1993). Gains higher than one imply an overestimation of the perceived inertial velocity with respect to the visual stimulus. This can stem from an over-registration of the vestibular signal in the brain, an under-registration of the visual velocity, or both. At the lowest velocities the gain is the highest. From visual experiments it is known that small gratings in the peripheral visual field can be underestimated (e.g. Johnston and Wright, 1986). The slower the motion, the stronger the underestimation. Such an effect may have occurred, although the solid angles of the visual stimuli in the set-up are rather large, see Figure 2. The high gains do not, however, occur in the tested yaw motion. It can therefore be concluded that the vestibular sensed velocity of the roll and swing motions are overestimated. The highest gains at the lowest velocities may, however, be explained by an underestimation of the visual

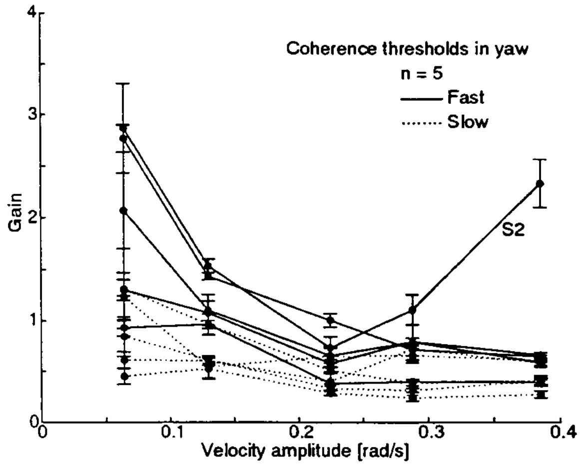


Fig. 11. Results from the yaw experiment. The error bar indicates twice the standard error of each threshold.

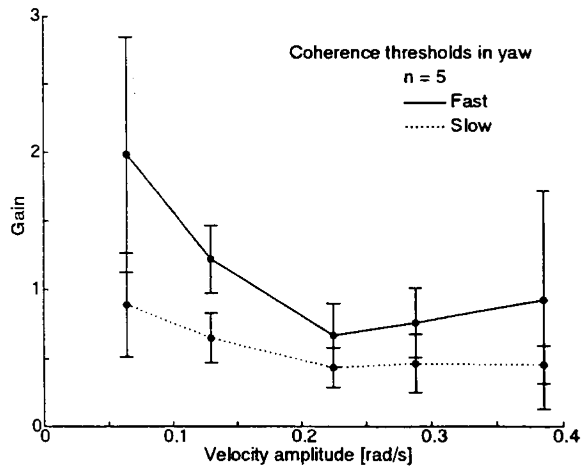


Fig. 12. Group result from the yaw experiment. The error bar indicates twice the group standard deviation.

velocity.

The gains for roll are higher than those for swing. There is an influence of the specific force, since the rotation component in roll and swing is identical for each motion tested. Apparently, the contribution of the specific force in the inertial motion, increases the self-motion sensation.

The lower thresholds and the narrow coherence zone in yaw may be explained by the increased familiarity of humans with motions of this type. The subjects can therefore register both the visual and vestibular component more precisely.

The results reveal differences among subjects. In roll and swing, subject 1 has higher gains than the rest of the group. To a lesser degree, subject 2 shows deviant gains at high velocities in yaw. Care has therefore to be taken while defining coherence zones in general since perceiving vestibular and visual velocity as coherent is an individual process.

5.2. *Implications for vehicle simulation*

Coherence zones are inherently related to self-motion perception. There exists some extent to which visual and vestibular stimuli may disagree while the observer retains a coherent perception. This is a useful phenomenon for the definition of stimulus generation in moving-base vehicle simulation. The major remaining obstacle in modern moving-base simulators is the narrow spectrum of inertial stimuli that can be generated. Motion filters are designed to translate motions occurring in the vehicle to motions of the simulator, in order to keep the excursion of the simulator within its limits. One of the main topics in motion filtering is the reduction of so-called false cues, or cues that remind the driver of the artificial situation. Such cues are fatal for realistic simulation. When stimulus generation stays within the coherence zone, however, no false cues will be perceived. The experiments described in this paper form first steps to obtain these zones. Direction coherence, and coherence thresholds at sustained visual stimulation, for example, are next steps to come to a more complete description of coherence zones.

Acknowledgements

This research is sponsored by the Netherlands organisation for scientific research (NWO). The National Aerospace Laboratory (NLR) is gratefully acknowledged for providing access to their 6-DOF research moving base.

6. REFERENCES

Hosman, R.J.A.W. and J.C. van der Vaart (1978). Vestibular models and thresholds of motion perception. *Delft University of Technology, Faculty of Aerospace Engineering, The Netherlands*, report LR-265.

Hosman, R.J.A.W. and F.A.M. van der Steen (1993). False cue detection thresholds in flight simulation. *Proceedings of the AIAA Flight Simulation Conference* 93-3578.

Howard, I.P. (1982). *Human Visual Orientation.* Sussex: John Wiley.

Johnston, A. and M.J. Wright (1986). Matching velocity in central and peripheral vision. *Vision Research* **26(7)**, 1099-1109.

Wertheim, A.H. and B.S. Mesland (1993). Movement perception during sinusoidal linear egomotion. *TNO Institute for Perception, Soesterberg, The Netherlands*, report IZF-1993-3.

HUMAN OPERATOR ADAPTATION TO A NEW VISUO-MANUAL RELATIONSHIP

Olivier Guédon, Jean-Louis Vercher and Gabriel M. Gauthier

Laboratoire de Contrôles Sensorimoteurs, Université de Provence & CNRS URA 1166
Avenue Escadrille Normandie Niemen. 13397 Marseille Cedex 20 (FRANCE)

Abstract: An experiment is presented in which observers were requested to closely follow the motion of a video-displayed target by driving a cursor through appropriate hand/arm motion. The relationship between the observer's actual arm motion and the arm motion as perceived through the cursor motion was suddenly changed. The resulting adaptive control was quantified in terms of spatio-temporal error and spatial gain. The results allow to determine, in a model of the visuo-manual tracking system, which element(s) handle(s) the plastic changes and how the plastic elements interact.

Keywords: Tracking Systems, Human factor, Man-Machine interaction, Adaptation, Teleoperation, Model

1. INTRODUCTION

When first exposed to a telemanipulation task, such as that of moving a cursor on a display screen through mouse displacements, or steering a robot arm through video display of the remote robot and its working environment by acting on a joystick, an observer must learn the metrical and directional relationship relating hand and cursor or robot displacements. Daily observations show that the motor performance of the operator increases as training progresses to finally reach an optimal level. If suddenly the visuo-manual relationship changes, the operator performance will first be strongly affected but as practice goes on, the operator will undoubtedly modify adaptively his behavior and recover normal motor performance. The operator learns the new spatial and temporal transformations between the motion of his own arm and the resulting motion of the cursor or the robot end effector. The observed adaptive changes are similar to the ones observed in response to an alteration of the normal visuo-manual relationship by optical systems such as lenses (Gauthier et al., 1979; Droulez and Cornilleau, 1986) or prisms (Welch et al., 1993). Behavioral physiologists define adaptation as the process by which the performance of the operator returns progressively towards normal in spite of the persistence of the alteration (Welch 1974).

1.1 Visuo-motor adaptation in teleoperation.

While a fairly large body of data is available regarding adaptive control in the visuo-manual system as a response to optical alterations, little is known regarding learning of the visuo-manual relationship in teleoperated tools and reaction of the operator to alteration affecting a learnt relationship. Besides, more complex problems arise in telemanipulation (telerobotics) where the controlled mobile static (metrical) and dynamic (inertial) characteristics, sensed by the observer, will combine with those of the steering device. One may also predict that the rate, overall amount and nature of the adaptation observed in response to visuo-manual alteration will depend on the nature (random or predictive) of the target trajectories.

1.2 Model of the visuo-manual tracking system.

The model used to interpret the data collected in this study is schematically illustrated in Fig. 1. As in Bullock et al.'s model (1993), a main path handles the spatial-to-motor and directional-to-rotation transformation needed to carry out a spatially defined trajectory. It also features, as in Hollerbach's model (1982) a classical visual feedback loop, and a secondary feedforward path holding an internal model of the target trajectory, as in Kawato et al.'s model (1988). With fast predictable targets the overall system is driven by the "internal target trajectory representation" (the physiologists call it the task internal model) element in a feedforward manner (with respect to the visual input), since the visual feedback does not operate on-line. A global error becomes available at the end of the trajectory through the comparison of the target and the actual cursor trajectories. When the normal visual manual relationship is altered as in the present study, changes must develop in the internal structure of the system in an adaptive way since the performance is slowly restored and persists when the visual feedback of the tracking cursor is suppressed. Where in the model do these plastic changes develop ?

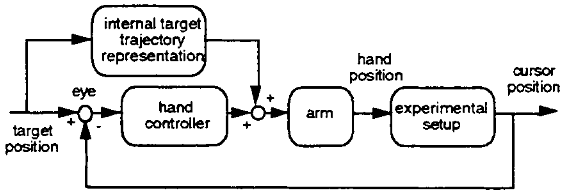


Figure 1: Model of the visuo-manual tracking system. Its peculiarity comes from the "internal target trajectory representation" element. For fast predictable targets the overall system is driven by this element in a feedforward manner (with respect to the visual input). Adaptive changes resulting from the alteration of the visuo-manual relationship may occur either/and in the "hand controller" and the "task internal representation".

The following section describes the adaptive changes which developed in observers submitted to sudden alteration of the relationship between their hand motion and the motion of a steered cursor on a screen. The tracking performance was evaluated in terms of spatio-temporal error and tracking gain along vertical (VT) and horizontal (HZ) directions. The protocol arrangements (tracking in visual closed- or open-loop conditions) and the nature of target trajectory (predictable or unpredictable) proposed to the observers during the tests and the exposure periods were designed to determine where, in the system, the plastic changes developed.

2. METHODS

The experimental setup (fig 2A) was designed to allow observers to hand track targets presented on a graphic screen. Vision of the arm was prevented. The target was a green dot while the tracking cursor was a white cross. The observer's 2-D hand motion was monitored by means of an infrared device whose sensitivity was adjusted to insure a one-to-one metrical relationship between hand motion and the resulting tracking cursor motion on the screen. First the observers (5 subjects were tested) practised with a one-to-one hand-to-cursor relationship (a 15-cm circular arm-stretched motion, in a vertical plan, resulted in a 15-cm circle on the screen). In the "closed-loop" condition, the observers were requested to accurately follow the target to minimize tracking error. In the "open-loop" condition, the cursor was not shown and the observer imagined its motion to follow the target.

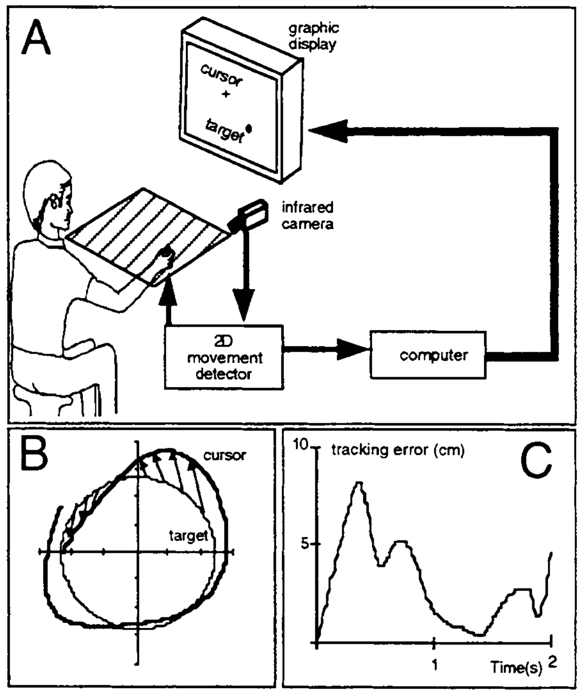


Figure 2: The observer was seated. facing a screen (A). Vision of the arm was prevented. The observer's hand steering motion (with the forelimb extended) occurred in a VT plane. The 2-D hand motion was monitored with an infrared device whose sensitivity insured a one-to-one metrical relationship between hand motion and the resulting cursor motion. Examples of tracking error vector on a trajectory (B) and time course during the trial (C).

To induce adaptive changes in particular sites of the visuo-manual tracking system and test these changes, two target motions were used. A 2-s circular trajectory was used as a fully predictable target motion. An unpredictable random-walk trajectory was obtained through the combination of sinewaves in HZ and VT directions. The shape and starting direction of the trajectory were randomly chosen by the computer. The same basic protocol was used in 4 successive experiments. The experiments started with 10 open-loop tracking trials (either circular or random trajectories), to determine the observers' "natural" visuo-manual relationship. Following this series, visual feedback was restored and the HZ visuo-manual relationship was suddenly increased by 2.5. Either 50 circles, as in experiment 1 and experiment 3, or 10 random-walk trajectories, as in experiment 2 and experiment 4 were then executed. The overall exposure time (100 s) was identical with 50 circles and 10 random-walk trajectories. At the end of this adaptation period, open-loop trials (circles or random trajectories) were presented to determine the effect of the training. Ten closed-loop circles or random trajectories was finally run with a normal visuo-manual relationship to follow the deadaptation process.

Experiment 1. The adaptation to the visuo-manual alteration was induced with a 2-s circular trajectory. The prediction was that such a target was likely to induce plastic changes in the "internal target trajectory representation" element and possibly in the "hand controller" element. The test also executed with the 2-s circular trajectory allowed one to determine the combined plastic changes occurring in these two structures.

Experiment 2. The adaptation was induced with a target moving randomly. The prediction was that the resulting adaptive changes were to develop only in the "hand controller" element. These changes were determined by testing the observers with random-walk trajectories.

Experiment 3. The adaptive target trajectory was a 2-s circle while the test target was a random trajectory. The prediction was that most of the adaptation would develop in the "internal target trajectory representation" element and could not be measured with random trajectories. However, if adaptive changes were measured with random trajectories they would have to be attributed to the "hand controller", meaning that predictable target still induce adaptive changes in the main direct pathway of the tracking system.

Experiment 4. The adaptive target trajectory was a random motion while the test target was a 2-s circular trajectory. The adaptive changes were expected to occur only in the "hand controller" element. The difference between the amount of adaptation between experiment 2 , and experiment 3 would be a measure of the cross adaptation between the "hand controller" and the "internal target trajectory representation" element.

VT and HZ components of target, hand and cursor trajectories, recorded at 100 samples/s, were saved on disk for off-line analysis. The subjects' (spatio-temporal) tracking error was computed as the distance between target and cursor positions at the same sampling time (Fig. 2B). This distance is nil only if cursor and target are at the same place at the same time. The tracking error over one trial was evaluated by computing the cumulated sum of the sampled errors along the whole trajectory. A global tracking error was defined as the total error along the trajectory divided by the length of the trajectory. The changes of tracking error, from trial to trial, were plotted as a function of trial order as in Fig. 4B and 6B. The adaptive changes were also described by the ratio, named the gain, between the size (HZ and VT extent) of the envelop containing the tracking response and the envelop containing the target trajectory. In open-loop trials, because the observer did not usually terminate his trajectory on the target end-point, the resulting HZ and/or VT "drift" was calculated and deleted before computing HZ and VT gains.

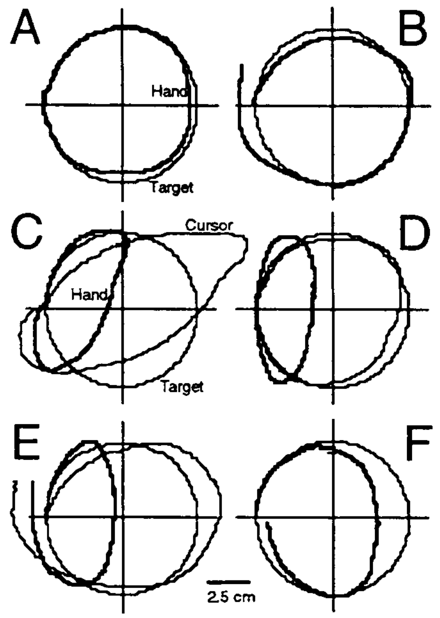


Figure 3: Selected tracking responses to illustrate adaptive changes. Superimposition of target, hand and hand-steered cursor trajectories emphasizes differences in the spatial tracking error between conditions with or without visual control of the cursor and before, during and after adaptive exposure.

3. RESULTS

3.1 Experiment 1.

In experiment 1, the observers adapted to, and were tested with circular trajectories. Figure 3A illustrates the tracking performance of an observer in the closed-loop condition before any alteration of the visuo-manual relationship. The VT and HZ hand tracking gains were fairly close to unity and the trajectory was smooth and circular. When visual control was not allowed, the response was still smooth and fairly circular but because on-line visual control of the cursor was absent, the response trajectory was sometimes shifted to the left (as in Fig. 3B) or to the right. When the alteration was applied to the HZ direction (X2.5 increase of the hand-to-cursor HZ sensitivity), the hand response in the initial 1 to 3 trials was markedly affected (Fig. 3C). During the first 20 trials, the hand motion trajectory progressively evolved to closely fit the target trajectory. Figure 3D illustrates the high hand accuracy recorded during one of the last exposure trials. The hand motion was then an ellipse with a ratio between VT and HZ axes equal to 2.5. The operator was fairly well adapted to the alteration, as shown in Fig. 3E, which applies to the first open-loop trial following the exposure. Indeed, the performance is close to the requested one in spite of the persistence of the alteration and the absence of visual feedback. The adaptive changes did not vanish immediately after returning to the original visuo-manual relationship. Indeed, Fig. 3F which illustrates the first trial in the closed-loop condition with a one-to-one hand to cursor relationship, definitely shows the effects of the plastic changes to the visuo-manual relationship. The hand motion trajectory was an ellipse, identical to the ones executed in the altered condition (instead of a circle). About three to four trials later, the hand tracking performance had returned to normal.

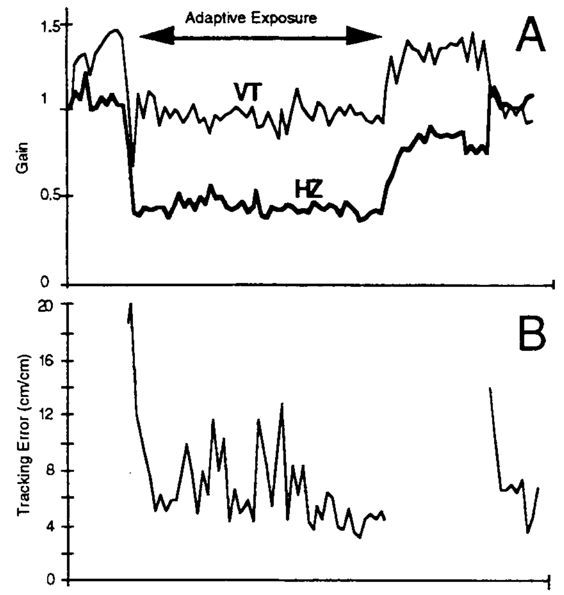


Figure 4: HZ and VT gain changes (A), as a function of trial order, resulting from a 2.5 increase in the hand-to-target sensitivity while tracking circular trajectories. During the closed-loop adaptive exposure the HZ gain decreased to match the alteration. When the normal condition was restored, the visuo-manual tracking performance recovered only after 2 to 3 trials. CL and OL refer to open-loop and closed-loop tracking. respectively. Time course of the tracking error (B) during the adaptive session illustrated in A. During the adaptive period the tracking error decreased rapidly to stabilize around a value close to the closed-loop error and normal hand-to-cursor sensitivity.

Figure 4A shows the time course of VT and HZ gain changes in one observer during an experimental session involving alteration of the HZ visuo-manual relationship. In the first open-loop series of trials, VT and HZ gains were slightly different from each other but still close to unity. During the all closed-loop adaptive exposure the tracking gains were, as expected, close to 1 and 0.4 for the VT and HZ directions, respectively. After the adaptive exposure the HZ gain was strongly reduced (approximately 0.6 for the first 2 trials) while the VT component of gain remained close to 1. As this open-loop sequence proceeded, the HZ gain increased continuously towards about 0.8. In the meantime, the VT gain returned slowly to its pre-exposure level.

The corresponding tracking error curve is shown in Fig. 4B. The average error in the first series of closed-loop trials was around 5 cm/cm. In the course of the adaptive exposure, the tracking error which was first large (around 20 cm/cm), decreased slowly towards around 5 cm/cm. The first trial following return to a normal visuo-manual relationship had a tracking error close to 15 cm/cm in spite of available visual control. Subsequently, the tracking error decreased rapidly towards pre-exposure level.

Table I: Normalized gains (5 subjects) in horizontal (HZ) and vertical (VT) direction after exposure to new relationship. M is the mean gain. Change is the M complement to 1, expressed in %.

Ss	Exp 1		Exp 2		Exp 3		Exp 4	
	HZ	VT	HZ	VT	HZ	VT	HZ	VT
1	0.47	0.93	0.73	1.13	0.85	0.88	0.67	0.79
2	0.53	0.84	0.62	0.92	0.87	1.00	0.52	0.83
3	0.72	0.85	0.63	1.00	0.90	0.86	0.79	0.80
4	0.55	1.00	0.71	0.96	0.86	0.94	0.94	0.82
5	0.54	0.86	0.63	0.84	0.66	0.78	0.71	1.13
M	0.56	0.89	0.66	0.97	0.73	0.87	0.83	0.89
change	44%	11%	34%	3%	27%	13%	17%	11%

Table 1 summarizes the changes of gain resulting from the period of adaptive exposure in the 5 observers and in the 4 experiments. The changes of gain in the VT and HZ directions appear as the difference between the normalized gain measured after and before the adaptive period. Normalization consists of referring the control gain (before adaptation in the non-altered condition) to 1. In experiment 1, the change of gain in the HZ direction after the exposure, averaged over the 5 observers, was 44% of the gain before the exposure. If full adaptation had occurred, a 60% change was expected.

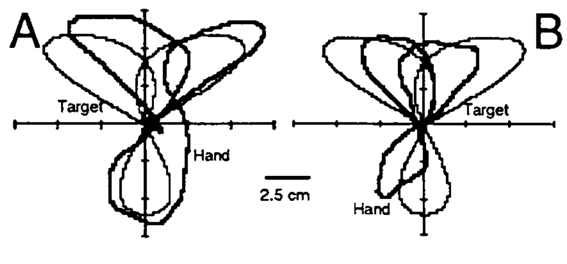


Figure 5: Tracking response from one observer before (A) and after (B) the adaptive period. Adaptation is evidenced by the decrease of HZ hand movement excursion between these two trials.

3.2 Experiment 2.

In experiment 2, the observers were tracking a randomly moving target both during the test trials and during the exposure period to visuo-manual relationship alteration. According to the model proposed in Fig. 1, the adaptive changes were expected to develop in the "hand controller" only. Figure 5 shows one observer's tracking response before (A) and after (B) the adaptive period. Adaptation is evidenced by the decrease of HZ hand movement excursion between these two trials. Fig.

6A illustrates the changes of VT and HZ gains in one observer during the entire experiment. In the first trial after the end of the exposure period, the HZ gain was 0.6 and subsequently increased slowly. The HZ and VT gain changes, averaged over the 5 observers, were 34% and 3%, respectively (Table 1). Overall, the change of HZ gain, resulting from the exposure period with random trajectories, was smaller than the change observed in experiment 1 with circular trajectories, where the tracking system was essentially driven through the "internal target trajectory representation" element. The tracking error curve (Fig. 6B) shows that during the adaptive exposure the global tracking error did not improve and rather remained to a level higher than in experiment 1.

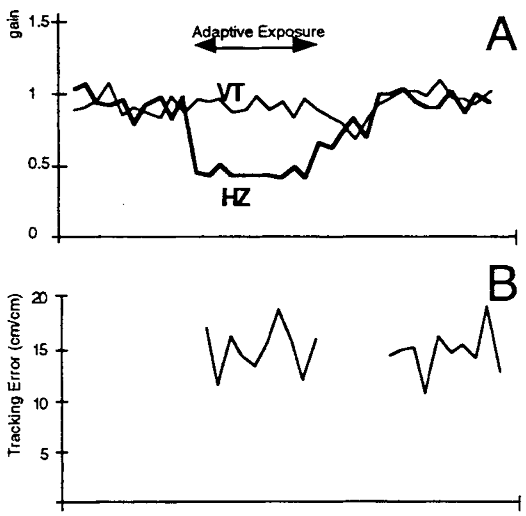


Figure 6: Changes of VT and HZ gains in one observer during the entire experiment. The time course of the adaptive changes are similar to the ones observed with a predictable target. In the first trial after the end of the exposure period. the HZ gain was 0.6 and subsequently increased slowly.

3.3 Experiment 3.

The observers were tracking 2-s circular trajectories during the exposure period and random motion trajectories during the gain tests. It follows that both the "internal task representation" and the hand controller" may be adaptively modified. The changes in the HZ and VT gains, averaged over the 5 observers, were 27% and 13%, respectively (Table 1).

3.4 Experiment 4.

The observers were exposed to random-walk targets during the exposure period and tested with circular trajectories. The hypothesis predicts that the plastic changes resulting from the visuo-manual alteration in such exposure conditions should only affect the hand controller element. The changes in the HZ and VT gains, averaged over the 5 observers, were only 17% and 11%, respectively (Table 1).

4. DISCUSSION

When the normal visual manual relationship was altered, adaptation must have developed in the internal structure of the system since the performance was slowly restored. This alone does not imply that long-term plastic changes developed. The visual feedback loop, assisted by the target trajectory internal representation could explain the performance recovery. However, since the changes of gain and tracking error were observed after exposure without visual feedback and in response to an input identical to that presented to the non-altered system, plastic changes must have definitely occurred in the arm tracking system to yield the appropriate output.

A global analysis of the data shows that the degree of adaptation reached after visuo-manual relationship alteration, as well as the time course of the adaptation process, were fairly different in the 4 experiments. All observers showed systematic adaptation with some degrees of variation in experiments 1-3 while limited adaptation developed in experiment 4. When adaptation occurred, there was a cross direction effect so that an adaptive decrease of gain of the HZ hand motion resulting from a sudden instrumental increase of sensitivity of the visuo-manual relationship in that direction, usually resulted in a concomitant decrease of gain in the VT direction as well. This phenomenon has already been described earlier (Bedford, 1994), though no explanation has yet been proposed to elucidate this point. As already mentioned in the introduction, visuo-manual adaptation in telemanipulation and hand pointing at visual target through optical systems (prisms or anisometric lenses) share characteristic features. In particular, the profile, the amount of adaptation reached after similar adaptive periods, and to some extent the adaptation and deadaptation time-course scales are similar. In telemanipulation, further aspects related to the shape and dynamics of the trajectory must be considered. For example, if the target path is predictable, the hand motor system will rely on predictive control (activation through the "target trajectory internal representation" path). If the target trajectory is a random motion, only the main feedforward and feedback paths of the visuo-manual tracking system will be activated. It follows that by comparing the amount of adaptation measured in various tracking conditions proposed to the observers during the exposure period and as a test for adaptation, one may determine where, in the system, the plastic changes develop when the visuo-manual relationship is suddenly altered (one assumes that most of the adaptive changes do not occur in the peripheral apparatus (the arm plant).

The model shown in Fig 7 is modified from the one in Fig. 1 to integrate the present data. It basically shows that when the target used during the exposure is fast and predictable (circle), the ensuing plastic changes do not develop only in the "internal target trajectory representation" but also in the "hand controller". The data also suggest that the "hand controller" is subdivided into two sub-elements: one handles the plastic changes occurring as a result of random target adaptation while the other handles the changes resulting only from exposure to unpredictable target trajectories, and is only activated with predictable target. The model arrangement is coherent with the changes observed and listed in Table 1.

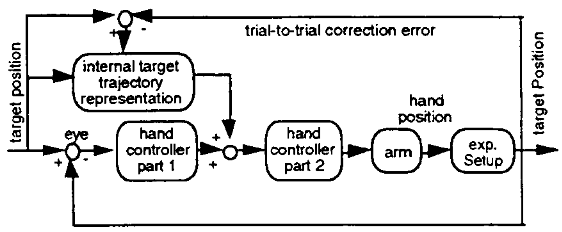


Figure 7: Improved version of the visuo-manual tracking system model which incorporates the data collected in the present study. A part of the plastic changes induced in predictable target tracking occurs in the "hand controller".

As a possible implication of the present report findings in telerobotics, let us consider a situation where the working environment is viewed first from a particular point, hence providing a given visuo-manual relationship to the interface linking the operator to the manipulated machine. If the view point suddenly comes from another camera situated closer to the operated machine (for higher resolution) or from a side-shifted camera (to observe another feature) the apparent gain will change. The first case is equivalent to a HZ and VT gain increase while in the second case, the gain change will be equivalent to altering the HZ gain while maintaining constant the VT gain. To compensate for this effect, the visuo-manual relationship could be automatically modified to maintain visuo-manual constancy, for example by changing joystick sensitivity with viewing camera distance. Indeed, Welch et al. (1993) have shown that repeatedly adapting and readapting to two mutually conflicting sensory environments leads to a separate adaptation to each condition, as well as an enhancement of the capacity to adapt to other alterations. This study may also contribute to better knowledge of visuo-manual control, particularly to the definition of the task internal model which, in spite of the efforts, over the past few years, (Bullock et al., 1993; Flash and Hogan, 1985; Moray, 1986; Kawato et al., 1988) has not yet provided an unified representation.

ACKNOWLEDGEMENTS

This work was supported by a grant from the Cogniscience/Cognisud CNRS program.

REFERENCES

Bedford, F.L. (1994). Of computer mice and man. Cahiers de Psychologie Cognitive/Current Psychology of Cognition, **13**, 405-526.

Bullock, D., S. Grossberg and F.H. Guenther (1993). A Self-Organizing neural model of motor equivalent reaching and tool use by a multijoint arm. J. Cognitive Neurosci., **5**, 408-435.

Droulez, J. and V. Cornilleau (1986). Adaptive changes in perceptual responses and visuo-manual coordination during exposure to visual metrical distortion. Vision Res., **26**, 1783-1792.

Flash, T. and N. Hogan (1985). The coordination of arm movements; an experimentally confirmed mathematical model. J. Neurosci., **5**, 1688-1703.

Gauthier, G.M., J.-M. Hofferer, W.F. Hoyt and L. Stark (1979). Visual-motor Adaptation. Quantitative demonstration in patients with posterior fossa involvement. Arch. Neurol., 36, 155-160.

Hollerbach, J.M. (1982). Computers, brain and the control of movement. TINS, **June**, 189-192.

Kawato, M., Y. Uno, M. Isobe and R. Suzuki (1988). Hierarchical neural network model for voluntary movement with application to robotics. IEEE Control Systems Magazine, **Apr.**, 8-15.

Moray, N. (1986). Modelling cognitive activities: human limitations in relation to computer aids. NATO ASI Series. Vol F21. In: Intelligent Decision Support in Process Environments (E. Hollnagel et al., Eds.), pp 273-291. Spring-Verlag, Berlin, Heidelberg.

Welch, R.B. (1974). Research on adaptation to rearranged vision: 1966-1974. Perception, **3**, 367-392.

Welch, R.B., B. Bridgeman, S.D. Anand and K.E. Browman (1993). Alternating prism exposure causes dual adaptation and generalization to a novel displacement. Perception & Psychophysics, **54**, 195-204.

GRAPHIC COMMUNICATION AND HUMAN ERRORS IN A VIBRATORY ENVIRONMENT

by L. DESOMBRE [1], N. MALVACHE [1] and J. F. VAUTIER [2]

[1] Laboratoire d'Automatique et de Mécanique Industrielles et Humaines
U.R.A. C.N.R.S. n° 1775 - Université de Valenciennes et du Hainaut-Cambrésis
Le Mont Houy - B.P. 311 - 59304 VALENCIENNES CEDEX-FRANCE
[2] Direction des Recherches Etudes et Techniques - Division Ergonomie
4 bis, rue de le Porte - d'Issy - 00460 PARIS - ARMEES

Abstract: The following study has been written in view of an ergonomical application. This paper is about an experiment on a simple decision making task. The purpose is to analyze human errors and to contribute to the cognitive modelisation of the human operator. Its originality is to try to take into account the influence of specific vibrations, in addition to the intrinsic characteristics of the human operator, .
After general considerations, the experimental protocol is detailed and some specific data processing are presented. The close examination of the results reveals a few characteristic behaviors, when the presented graphs are bar-graphs or polygons.

Keywords: Evaluation methods, graphical information, response times and performances.

1. PROBLEMATICS AND METHODOLOGY

In spite of the wide usage of visualization screens, the behavior of human operators remains an essential element in the reliability of high risk technological systems; such as the automatic drive of industrial processes or air, sea and land piloting. In the systems where the conversational and graphic methods may facilitate some of the operator's thought processes, the effects of decision and human actions are mostly transmitted and accentuated by automatism chains of control and supervision. In such conditions errors could have serious repercussions which could invite the scientific and military communities to examine the nature of human error in order to develop new solutions to increase the reliability of the shift in systems from man to machine (Reason, 1993; Masson, 92; Leplat, 85).

The willingness to deal with human error is found in this study where the operator is subjected to visual-postural disturbances (Desombre, 93) that may affect their "functioning" during a simple-natured repetitive task that call for visual perception and where the complexity lies in the part of thought necessary to:

- The classification of graphic information according to several criteri defined in advance,
- To the feelings of certainty and difficulty expressed by the operator regarding the information captured.

In attempting to clarify the respective roles of information and disturbances in the organization of the "visual-postural performances," the analysis of the results through the discovery of important indicators inherent in our reasoning, will perhaps assist toward a better understanding of the operator's behavior controlling a process.

In all, there is a situation of a reliable and ergonomic nature favorable to the study of the relationship between the processing of visual-postural information and the mental representation (operative) linked to this processing through objective and subjective measures.

2. DESCRIPTION OF THE EXPERIMENT

In order to facilitate the understanding of the results shown in paragraph 3, you will find a detailed description of the experimental device below.

2.1 Principle of the experiment

The set-up in the work studied is of an operator who in front of a screen who identifies a geometric shape belonging to a defined class and detects the presence of a possible graphic abnormality.

During his work, the subject is under static and vibrational states. The responses, as well as the different response times, are recorded permitting an analysis of the errors. They also offer valuable indications of the control of the processes unobservable that imply human behavior: motor-sensory activities of ocular exploration and cognitive activity.

2.2 Choice of context

The experiment takes place in an isolated room where the human operator is subjected to a disturbed visual-postural environment.

a-) Stimuli

Figure 1 shows the geometrical shapes used. Only 2 types of graphics can be distinguished: the bar-graphs and polygons with n=6,8 and 10 bars or vertices.

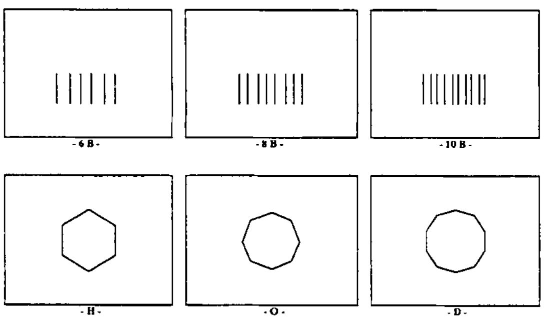


Fig. 1. Presentation of the utilized but undamaged geometrical shapes named with letters that will be used as references.

The choices linked to the duration of the experiments are related to 2 types of graphics in use. Nevertheless it is quite obvious that the 2 supports of the presentation of information will need the help of various functional operations: especially in the area of the visual exploration.

From a methodological point of view, standardization between the study made with the bar-graphs and the polygons had been researched to facilitate the comparison of the results.

b-) Experimental platform

The works performed relies on an already existing mechanical structure (existing at the L.A.M.I.H). It is based on the principle on which the gyroscope works so as ensure an uncoupling between the three axes: roll, pitching and twist.

The vibrations retained for this study are generated by a movement of 1 Hertz sinusoidal rolling with 10° amplitude peak to peak (figure 2).

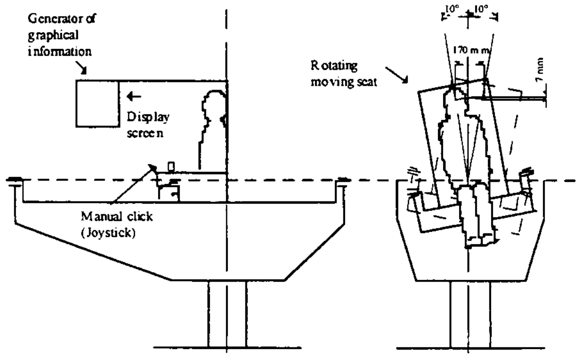


Fig. 2. Diagram of the experimental platform.

The monitor or the generator of graphical information is firmly attached to the moving of the chair on which the pilot is seated, and is controlled by a joystick and a manual button.

c-) Graphic abnormality

Concerning the display of abnormalities, the deteriorations of the stimuli are introduced, in a random way, in the series of graphics presented to the human operator. An example of graphical deterioration is presented in figure 3 with n=6.

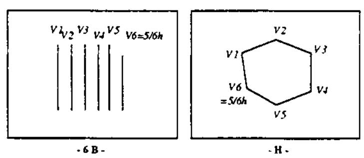


Fig. 3. Example of graphics (6B, H) with anomalies - part anomaly V6=5/6h, h corresponding with the variable height of the bars or with the magnitude of the diagonal of the polygons.

2.3 Subjects

A group of 27 subjects has been tested during the research period. They are divided in 3 groups:

- 4 subjects during the preliminary phase so as to precise and improve the tests or experiment units,
- 10 subjects during the first series of experiments (I) taking into account the whole range of the parameters of the study,
- 13 subjects tested for the last experiment (II) in order to confirm or to invalidate some interesting points concerning the preceding study.

The average age of the subjects is 22 (years old). The subjects are all French speaking science graduate students. All of them were in good health and all used their dominant hand in the execution of their responses.

2.4 Experimental task

After a series of preliminary tests and considering our results (Roger, 1985), a succession of 3 levels of difficulty with N=5,7,9 classes has been established in order to obtain a reference which may, afterwards, be consolidated by extending the N level or difficulty and by using other graphical forms.

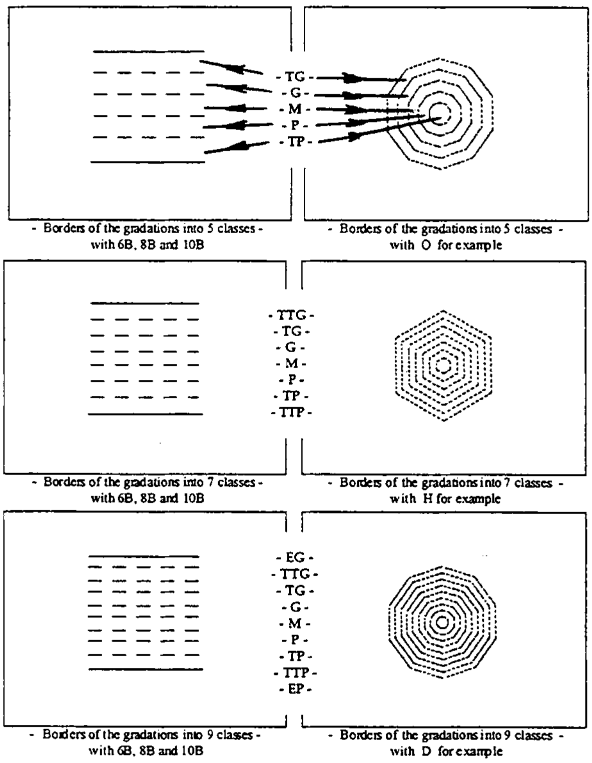


Fig. 4. Linguistic expressions defining gradations so as to classify into 5,7 and 9 classes (EL=Extremely Large, VVL=Very Very Large, VL=Very Large, L=Large, M=Medium, S=Small VS=Very Small, VVS=Very Very Small, ES=Extremely Small.

In the conditions described above, the task of the human operators is to classify the graphics (fig 1) according to their size. This classification is made possible by identifying the visual shape in one of the unconnected N subsets (fig 4) defined by learning and sharing the possible sizes. The subject must also detect the presence of possible graphic abnormalities (fig 3) in the visual shapes he has not learned in his training period.

After his classification of the graphical size in the corresponding class, the subject subjectively evaluate the certainty of his choice and the difficulty by positioning on the screen the cross on a line segment, graduated at its ends by 0 and 1 and controlled by a joystick (fig 5).

Bearing in mind that most activities imply decisions which involve risk, the processing of the uncertain inherent knowledge and our faculty of reasoning with the help of subjective indicators (Desombre, 93), has made an interpretation of the possible errors and omissions .

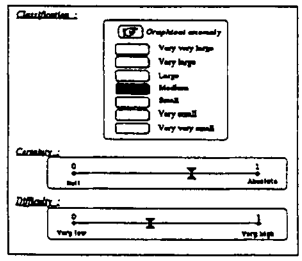


Fig. 5. Example of classification of a medium graphic and of subjective evaluations of a human operator (the size has been predefined into 7 gradations: Very very small,..., Very very large).

2.5 Experimental procedure

Each subject devotes 3 mornings to the experiments during which a single level of difficulty N is tackled. Each morning is composed of the 6 experimental units presented in figure 6, each one preceded by a learning phase of the borders of the gradations of the classes and tests, containing arbitrarily of 70 stimuli.

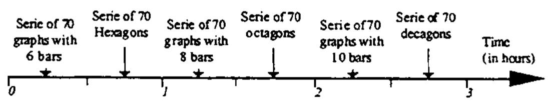


Fig. 6. Experimental morning for the subject comprising 6 experimental units: 3 involving bar-graphs and 3 involving polygons, with 6,8,10 bars or sides.

The subject starts with the classification into 5 classes (first morning), then 7 (second morning) and ends with 9 classes (third morning).

The experimental units are done in a postural stable position disturbed by phases during which the subject is submitted to vibrations, example fig 7.

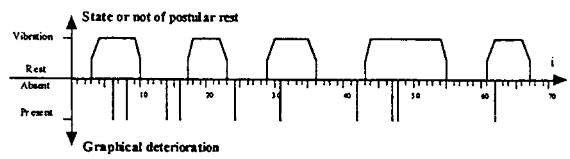


Fig. 7. Sequence of graphical stimulus i, with or without graphical anomalies, during or not postural disturbances.

The classification sequences submitted or not to rolling vibrations are equally distributed during the whole time of the experiment (figure 7). The subject carries out his task with no time limit. Time is registered for the sole purpose of an analysis.

2.6 Recording the data

The output variables are the response-times the subject takes to decide as to the classification of the graphic and degrees of certainty and difficulty as well as their corresponding values, (figure 8).

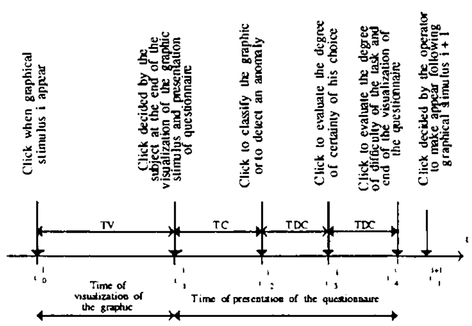


Fig. 8. Sequence of presentation of graphical stimulus and of the human operator's making decision.

The output variables are the following ones:

• decision time

i = number of the i stimuli = [1,70],

$TV = t'_1 - t'_0$ = time to display the graphic,

$TC = t'_2 - t'_1$ = classification time,

$TDC = t'_3 - t'_2$ = time to evaluate the degree of certainty,

$TDD = t'_4 - t'_3$ = time to evaluate the degree of difficulty.

• Choice of gradation

C_N = choice of the class gradation [Small, medium,..].

• Value of degrees of certainty and difficulty

DC = value, ranging from 0 to 1, of the degree of certainty,

DD = value, ranging from 0 to 1, of the degree of difficulty attributed by the subject.

3. RESULTS

The results obtained of the main experiment I and are confirmed those stemming from experiment II one year later. The discussion will be mostly based on the quantitative results (validated or not through the Wilcoxon Test), and as well as on the data obtained through the observation of "cumulative membership function".

No meaningful difference was noticed either between the various times of evaluation of the degrees of certainty and difficulty or between the various of n=6,8,10 regarding the output variables. Only a slight difference between n=6 and n=8, or 10 was noticed, rather insignificant on the threshold of .05.

3.1 Influence of the shape exposed

a-) On the performance of the subject

The operator's performance is defined by the percentage of right responses obtained during the classification and the detection, at the time of the sequence. An estimation of the rate of error of the subjects as a whole can be made according to this indication (figure 9.A).

The performances obtained with the bar-graphs are higher than those obtained with the polygons with N=5,7 classes (this difference is rather significant with the threshold of .01 in 5 classes, and .005 in 7 classes). There is no difference when N=9 classes. Furthermore, the rate of errors increases for the bar-graphs during the transfer from N=7 to N=9 ($p<.0003$).

b-) On the visualization time

The total values shown in figure 9 manifest clearly prominence to a longer time of visualization with the bar-graph (when compared with the polygons). This results is confirmed through tests of validity, which underscore that this variable is significant, no matter what the numbers n or N are. The inversion of classes between N=7 and N=9 also shows a significant effect on the errors. This tends to susbtantiate the idea of a reduction of the visualization time.

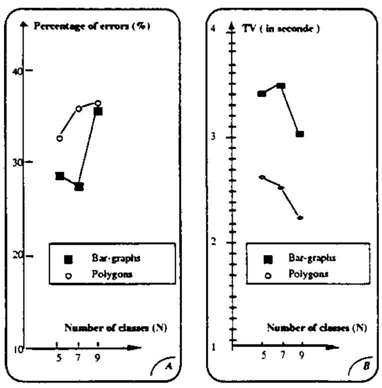


Fig. 9. Average value of the errors (A) and of the times of visualization (B) classified, for 9 subjects.

3.2 Linguistic responses, and overlapping zones

a-) Method

The method used in the analysis enables to notice the subjective aspect of the human operator. It could be defined as a creation of a "cumulative membership function" for every defined class (Malvache, et al., 1989). The function represents normalized summation of occurrences of the concerned adjective versus the dimension of the graphical stimulus.

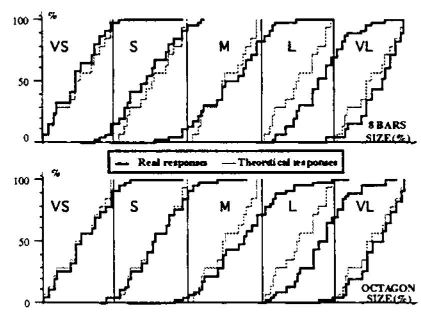


Fig. 10. Cumulative membership functions from the theoretical and real classification of the bar-graphs and polygons, presented below with n=8 and N=5 as examples.

Those functions enable to reveal the existence of overlapping zones between classes and in the internal model of the human operator (example fig. 10). theses areas are boundaries between classes and for the operator. It is appropriate to analyze those overlapping zones because errors always occur at that level.

The following method has been processed. Considering 2 cumulative membership function x and y, having a overlapping in the interval between sizes [a,b], function f(x,y)=min(x,1-y) is calculated. This function somewhat represents a function of distribution between x and y, at its bordaries.

b-) Result

The perceptive and mnemonic difficulties encountered by the subjects in classification of graphic stimuli can be precised with this processing. The analysis of the overlapping zones leads to a distinction of the identifiable errors between 2 classes, where the overlapping does not exactly adjust with the bordaries of those 2 classes.

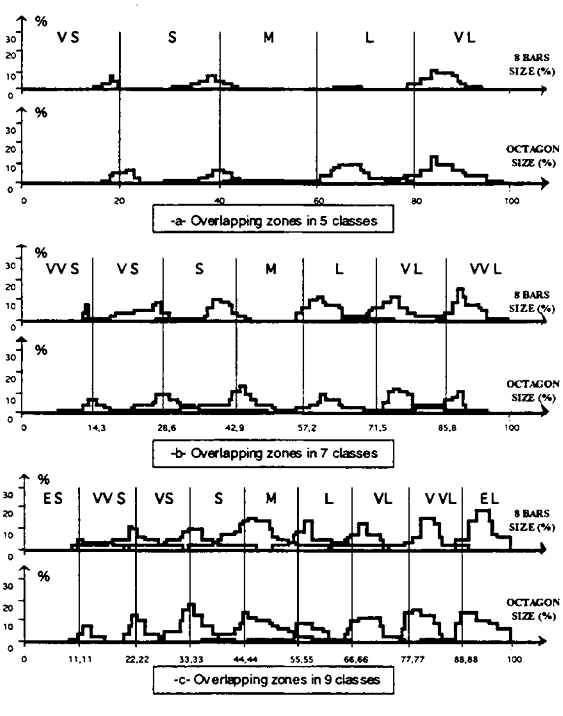


Fig. 11. Overlapping zones stemming from the valuable classification of bar-graphs and polygons, here with n=8 and N=5,7,9 as examples.

The figure 11 presents shift to the left or to the right of these overlapping zones. The tendency of the subjects in 5 and 7 classes and with the bar-graphs, to under- or overestimate the height of the bars depending, on the average axle of the glance (in the average classification) is then made obvious. On the contrary, in 9 classes, or with the polygons, the subjects always tend to overestimate the dimension of the graphic shape exposed.

The second experiment II made with the sole use of the bar-graphs and without any vibrations give the same results, once again.

3.3 Vibrations influence on the visuo-postural performances

The examination of figure 12 clearly indicates that the percentage of errors in a vibratory system is higher than the percentage of errors in a stable system regarding the polygons (p<.0001 for N=5; p<.0007 for N=7; p<.02 for N=9) and regarding the bar-graphs in 9 classes (p<.03 pour N=9). Those results are very similar to those previously defined, and enable to notice the perceptive changes with bar-graphs in 9 classes.

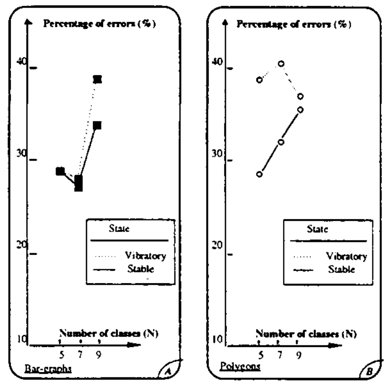


Fig. 12. Average values of the errors and for 9 subjects for both the bar-graphs and the polygons, in two different states: vibratory and stable.

Thus, considering the times of visualization and classification, figure 13, the following facts can be noticed:
- bar-graphs: time of visualization in a vibratory system is shorter in a vibratory state than in a stable state,
- polygons: insignificant effect.

The subjects don't need as much time to visualize the bar-graphs when submitted to a vibratory state.

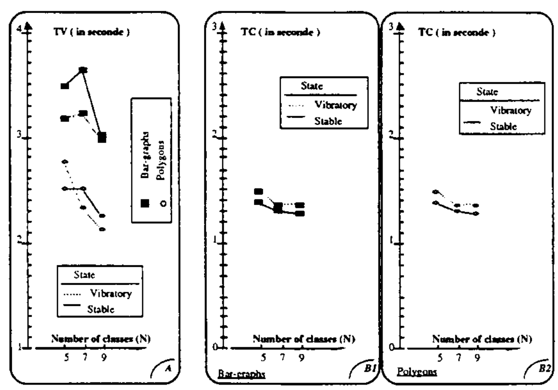


Fig. 13. Average values of times of visualization and classification for 9 subjects for bar-graphs and polygons in both states (stable and vibratory).

Figures 13 B-C also notice that the time needed for the classification under a vibratory state is higher than the time of classification under a stable state, for the bar-graphs (p<.0002) and for the polygons (p<.0019). This result is certainty linked to the motor-sensory activities of joystick.

4. DISCUSSION AND CONCLUSION

All the data examined allows us to stress several facts.

- Firstly, the performances generally obtained during the experiments with the bar-graphs are higher than those obtained with the polygons. This confirms the research in experimental psychology made from the graphic shapes containing several perceptive and variable dimension (height, shape, contrast,...).

In the following study, the bar-graphs (bearing in mind that a variation in height leads to a variation of shape), have a supplementary variable (the shape) compared to the polygons which always keep the same shape regardless of the variation of the relation between height/width.

This corroborates the majority of the data from the scientific literature which specified the conditions in which the addition of one or several dimensions to an existing dimension for a graphic shape leads to an increase in performances in a discriminative task (Garner, 1978; Coury, et al., 1988).

- The accentuation of the two perceptive data entry mechanisms emerges from the overall results concerning the overlapping zones and visual-postural performances. These results are very important because they show that data entry and the processing of information in bar-graph format, are done according to the complexity of the task (between 7 and 9 classes) and generally present the same characteristics as the polygons.

Thus, the classification process and its structural representation aid in finding an appropriate method of storage (ie. in relation to the average axe of the eye) and the reprocessing of the information by computer.

- As far as the temporal administration of the graphic messages is concerned, we notice a decrease in the visualization time in a vibratory state when compared to a stable state, when the bar-graphs are the stimuli. Thus, the bar-graphs are modulated through the postural state of the subject. We can see the importance of the time variable once again, when the perceptive mechanisms become more complex, the operators decrease the duration of the messages.

This result suggests the motivation to use a technological visual support according to the level of vibration the pilot is submitted to.

By way of conclusion, we know that graphics can facilitate a certain understanding of the operator and reinforce its reliability. Nevertheless, a certain amount of caution is necessary. One must not substantiate without reserve, the technicians' current tendency to make the screens out to be a privileged medium of visual information.

The study of the different data entry mechanisms of perception and processing of graphic information represents an intermediary indicator necessary to the understanding of the operator's responses to the type of work environment provided. The result are far from being perfectly codified and standardized but they can already be used in short term applications and in concrete situations

ACKNOWLEDGMENTS

Support for this experiment was provided by contract n° 90/204 convention n° 90.34`.204.00.470.75.01, through the L.A.I.H. (Laboratoire d'Automatique Industrielle et Humaine), from the D.R.E.T. (Direction des Recherches et des Etudes Techniques; Ministère de la Défense, Division Ergonomie Sciences Humaines).

REFERENCES

Coury, B.G. and J. Purcell (1988). The bargraph as a configural and a separable display. (Tech. Report CR-39-89/CTR DAAA 15-85-K-0010). University of Massachusetts, Amherst, Dept. of Industrial Engineering and Operation Research.

Desombre, L. and N. Malvache (1993). Reliability and visual workload of a pilot performing a simple decision making task and submitted to very low frequency vibrations. Proceedings of *XII. European Annual Conference on Human Decision Making and Manual Control*, pp. 2-19 - 2-25. June 22-24, 1993, Kassel, Germany.

Garner, W.R. (1978). Aspects of a stimulus: features, dimensions, and configurations. In E. *Rosch and B.B. Lloyd (Edt.), Cognition and Categorisation.* Hillsdale, NJ. Erlbaum.

Leplat, J. (1985). Erreurs humaines et fiabilité humaine dans le travail. Ed. A. COLIN, Coll. U, Paris, 1985.

Malvache , N. and D. Roger (1989). Knowledge extraction by variation analysis of the inner model in a simple decision making task. Proceedings of *4th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Man-Machine Systems.* September 12-14, 1989, XI'AN, China.

Reason, J.T. (1993). *Human Error.* Cambridge: Cambridge University Press.

Roger, D., C. Binot, G. Beyaert and D. Willaeys (1985). Man-machine dialogue and graphic communication: the problem of valuation. Proceedings of *V. European Annual Conference on Human Decision Making and Manual Control*, pp. 145-154. June 3-4-5, 1985, Berlin, Germany.

APIIS: A METHOD FOR ANALYSIS AND PROTOTYPING INTERACTION INTENSE SOFTWARE

Lúcia Vilela Leite Filgueiras
Selma Shin Shimizu Melnikoff

Escola Politécnica, Universidade de São Paulo, Brazil
Department of Computer and Digital Systems Engineering

Abstract: This paper presents a method to systematized requirement analysis and prototype design of interaction-driven software. The method is based upon extraction of objects and its attributes from a task-analysis description. The main goal is to capture the operator's mental model and map it into an object-oriented human-computer interface that can assure appropriate operation performance.

Keywords:human-machine interface, object modelling techniques, software engineering, human-centered design, operators, tasks.

1. INTRODUCTION

This method helps the human-computer interface analysis and prototyping of interaction-intense software. Interaction-intense software can be defined as application software where operator action is relevant to the performance of the system. In these cases, system actions are driven by operator inputs, as in transaction processing. Operation of process control systems is also another example of such systems.

Today, there are many software tools to build human-computer interface. Most of them have a simple dialogue: form-filling and graphical composition turns human-computer interface into a task accessible for non-experts in computers techniques.

This ease-of-use does not help, however, designing human-computer interface or even the screen layout for appropriate operation - this depends on the comprehension of user mental model about the process. Reproducing the user's mental model is even more important when the computer system comes to help or substitute some manual-performed activity.

APIIS, which stands for Analysis and Prototyping Interaction-Intense Software, intends to build a object-oriented model from a task analysis procedure. The operational procedure approach intends to reduce navigation through windows. This method has proven to be useful also when configuring human-computer interface software for process control.

APIIS is based upon object-oriented analysis and design, which has proven to be an efficient tool to map real world into software. Human-computer interface requires a design for consistency that impose that all objects are considered at a time, to assure objects and actions ortogonality.

The following phases are required to design the human-computer interface:

- definition of operational sequences;
- refinement of operational sequences in a detailed description script;
- identification of classes, objects, their attributes and methods;
- syntactical design of each identified object;

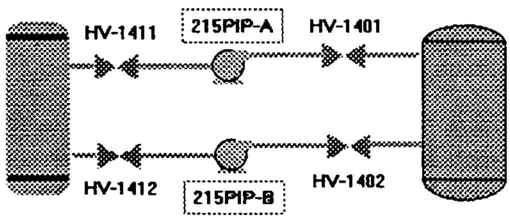


Fig. 1. Polymer injection system

- visual implementation of each object faceplate and its behaviour;
- window design to support operational sequences;
- evaluation of operational appropriateness

2. APIIS

In order to illustrate the method, this paper presents an example extracted from an continuous process control system. The operation chosen is the injection of a polimer into a vessell, illustrated in Figure 1.

2.1. Definition of operational sequences

This step is responsible for requirement analysis of the interaction-intense software.The designer must identify the routine tasks performed by the operators, and also identify exceptions to those tasks. Identifying all tasks in a process is a hard work, and even more difficult for computer systems designer who are seldom experts in the application field.

APIIS recomend the documentation of important operational procedures in a event graph, as the one shown in figure 2. The correct procedure (bubbles at left) and the exception to the routine (bubbles at right) can be stated. In computer-based dialogues error recovery is of utmost importance and must be stated as soon as possible.

2. 2. Refinement of operational sequences

Operational sequences identified as "bubbles" in the preceding step must be detailed in a description script. This script will show the information requirements to comprise the described activity.

The Operational Procedure Script (OPS) used is a mix of programming language and task analysis forms suggested by Burgy (1983) and Carey and Whalley (1990).

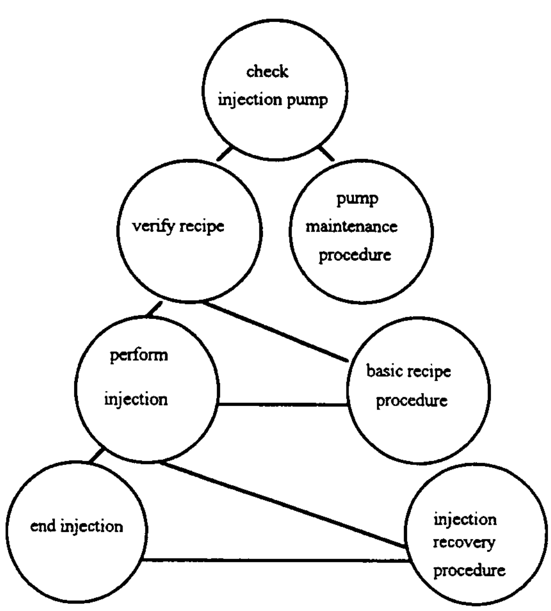


Fig.2 - Event graph for polymer injection

Task analysis is a well-defined procedure used in operational data gathering, which has proven to be useful in determining precisely operational functions, mental load evaluation and analysing "a posteriori" operational incidents. However, forms used in task documentation lack repetition and decision structures common to computer languages.

The task detailing considers the use of well-defined verbs as those from the Berliner taxonomy of operator actions (Dougherty, 1988).

Some other verbs may be necessary to cover human-computer dialogue, such as "point" and "click", but it is suggested that no verb but those defined be used in documenting procedure scripts, for they have a strict meaning.

OPS layout and refinement of task sequences identified in the previous step of the example are shown completely in Table 1.

Observe that most sentences are constructed from a subject-verb-object-attribute syntax. Some cases are a little more complex and may require some more lines of definition.

"Calculate", for instance, requires a multi-parameter function, that is many object-attribute pairs, though the calculation algorithm may be not important to the human-machine interface.

Decision and repetition structures can be added to simplify understandin non-linear tasks.

Table 1 OPS for polymer injection sequence

	subject	*action/complement*	*object*	*attribute*	*value*	*exception*	*exception routine*
	OP1	identifies	Injection pumps 215PIPA and 215PIPB	Maintenance status	on-service-	no pumps available	pump maintenance procedure
	OP1	decides	Injection pump	Maintenance status	out-of-service		
		based on	Injection pump	Recomended max. time of operation			
		and	Injection pump	Total time on service			
	OP1	sets	Injection pump	Maintenance status	out-of-service		
	OP1	checks	Recipe	Volume of injection		inventory not enough	basic recipe procedure
	OP1	calculates	Injection Pump	time on		wrong calculus	repeat step
		based on	Recipe	Volume of Injection			
	OP1	adjusts	Timer	time-on		wrong value	set new number
if Injection Pump active is 215PIPA							
then	OP1	sets	HV1401	operating status	open	valve does not open	switch to backup
	OP1	sets	HV1411	operating status	open	valve does not open	switch to backup
	OP1	sets	215PIPA	operating status	on	pump does not turn on	switch to backup
else	OP1	sets	HV1402	operating status	open	valve does not open	switch to backup
	OP1	sets	HV1412	operating status	open	valve does not open	switch to backup
	OP1	sets	215PIPB	operating status	on	pump does not turn on	switch to backup
end							
	OP1	sets	Timer	time remaining	start		
	OP1	monitors	Timer	time remaining	> 0	timer does not respond	stop sequence
	OP1	monitors	Vessel	level value		level high	stop sequence
	OP1	detects	Timer	time remaining	=0	failure to detect time end	injection recovery proc.
if Injection Pump active is 215PIPA							
then	OP1	sets	HV1401	operating status	closed	valve does not close	next step
	OP1	sets	HV1411	operating status	closed	valve does not close	next step
	OP1	sets	215PIPA	operating status	off	pump does not turn off	injection recovery procedure
else	OP1	sets	HV1402	operating status	closed	valve does not close	next step
	OP1	sets	HV1412	operating status	closed	valve does not close	next step
	OP1	sets	215PIPB	operating status	off	pump does not turn off	injection recovery proc.
end							
	OP1	informs	Supervisor	Conclusion of task			

2.3. Object-oriented modelling

Operation description scripts are used to identify classes, objects, their attributes and methods.

Careful detailing of operational sequences often leads to the identification of aditional, non-conventional human-computer interface features that contributes to a performance optimization and task time reduction. In the example, the object "pump" has many attributes which would not be defined if the pump was considered only as the installed equipment. In an application of APIIS method, it was observed that 60% of time consumed in a certain task was due to search of written documentation. Automation of this task was rather simple and at a low cost, and resulted in large benefit of average task time.

The transcription of OPS into object-oriented representation is now described.

The following definitions are necessary:

- *observable attributes*: these are object attributes that contains necessary information for the user to perform some task. The origin of information can be diverse: they can be collected from the process instrumentation or they can be generated by automatic control. Observable attributes are information that must be perceived by the operator; thus, they are associated to the perception and cognitive tasks.

- *controllable attributes* - these are object characteristics that must be modified by the user in some task. Because of importance of feedback in a interaction-intense software, all controllable attributes are also observable by default.

Notice that observable atributes can be *dynamic*, that is, subject to change like the status of a pump or *static*, if they always keep the same value - for instance the engineering units of a analog variable. Controllable atributes are always dynamic.

Objects are extracted from the columns "object" of OPS forms. For each object, the fields "parameter" and "value" inform about the attributes. If the Berliner verb is one of a sensorial process or a cognitive process, then the attribute is observable. If it is associated to a motor verb, then it is a controllable attribute.

Values for attributes can also be collected from the OPS description. Observe that motor tasks embed a decision for automation. However, notice that a manual backup may be necessary. The OPS form can help in determining the balance between automation and human action.

In a proces control application, in which the plant equipment and process are defined in terms of *instances*, OPS deals with those instances, not classes. It is the designer's work, then, to identify classes through the analysis of common attributes.

An example is the class "valve"- it is an "on-off" valve; the analysis of other OPS for the same plant could have revealed other classes.

The class definition is important to demonstrate consistency among operations, for the service defined for a parent class is inherited by a child class. The syntactic definition of this service should be applied automatically to all inherited classes.

Figure 3 illustrates the object model extracted from the example given. Observe that methods and alarm attributes can also be extracted.

2.4. Object design

Object design corresponds to the definition of syntactical behaviour of objects, that is, it defines objects faceplate and identifies graphical correspondence of attributes. Object design can be done after analysis of the main OPS describing the operational behaviour.

Object window is composed by:

- static description formed by the bitmap at the background plus references to static observable attributes;

- dynamic description, which can be associated to observable dynamic attributes and

- sensible field description, correspondent to controllable attributes.

Implementation of each object faceplate and its behaviour can be done using one of the many tools for human-computer interface design. Unfortunately, most human-computer interface softwares for process control are still not object-oriented, and this represent a great difficulty in design.

2.5. Display design

Display design for supporting each defined operational sequence can be comprised by the analysis of the corresponding OPS.

Object: Injection Pump (instances: 215PIPA and 215PIPB)

Attributes:	*Classification:*
Maintenance Status	controllable
Recomended maximum time of operation	observable, static
Total time on-service	observable, dynamic
Time-on	observable
operational status	controllable

Methods: Set Maintenance
Set Operational Status Open/Close
Alarms: No pump available
Pump does not respond to command

Object: Recipe

Attributes:	*Classification:*
volume of injection	observable, static

Methods: none
Alarms: none

Object: Timer

Attributes:	*Classification:*
time-on	controllable
time remaining	controllable

Methods: Adjust timer
Start
Alarms: End of timing

Object: Valve (instances: HV-1401, HV-1402, HV-1411, HV-1412)

Attributes:	*Classification:*
operating status	controllable

Methods: set operating status open/close
Alarms: valve does not correspond to command

Object: Vessel

Attributes:	*Classification:*
level value	observable, dynamic

Methods: none
Alarms: level too high

Fig. 3. Object model extracted from the OPS forms of the example.

This can assure that every attribute needed to cope with the task is present in the same screen, reducing the need for navigation, which is the main (or maybe the only) drawback of computer interfaces for complex systems.

When designing displays, one must go back to the instance definition in OPS.

2.6. Evaluation

After implementation of the prototype, operational appropriateness evaluation starts with the execution of primary sequences.

An efficient measure of appropriate information grouping can be obtained by "electronic link analysis", which is the computing of how often an operator visit some window when executing a given task.

3. CONCLUSIONS

This paper presented APIIS, a method for systematically analyse and prototype human-computer interface for interactive intense software.

APIIS is very useful when there is a manual activity to be analysed. This is often the case of automation applications, specially when associated with human-computer interface configuration software.

The method could be applied also if one is devising the system. In this case, the designer has to provide a task analysis as it "should-be" and be careful when evaluating operational appropriateness.

Experiments realized with this method have shown that it is efficient in obtaining the information requirements for completion of each task.

Incorporating time requirements in OPS is interesting to guarantee that the action can be performed in required time.

Task description can be complex depending on the nature of the cognitive process involved. Analysis must also consider operation that requires interpersonal cooperation and concurrent operation, which requires parallel information processing by the operator.

Presently, researchers at Escola Politecnica are presently working in the automation of this method.

4. REFERENCES

Burgy, D. *et al.* (1983). *NUREG/CR-3371: Task Analysis of Nuclear Power Plant Control Room Crews*; Nuclear Regulatory Commission, U.S.A.

Carey, M and S. Whalley (1990). Cognitive Task Analysis Techniques in the Design and Evaluation of Complex Technological Systems. In *11th. Advances in Reliability Technology Symposium.* (P. Comer (Ed.)); Elsevier Applied Science.

Dougherty Jr., E. M. and J. R. Fragola (1988). *Human Reliability Analysis - A Systems Engineering Approach with Nuclear Power Plant Applications.* John Wiley and Sons.

INTEGRATION OF COGNITIVE ERGONOMICS CONCEPTS IN KNOWLEDGE BASED SYSTEM DEVELOPMENT METHODOLOGIES

Michèle DURIBREUX, Christophe KOLSKI and Bernard HOURIEZ

L.A.M.I.H., University of Valenciennes
B.P. 311 - 59304 Valenciennes, FRANCE
e-mail: duribreux@univ-valenciennes.fr

Abstract: A Knowledge Based System (KBS) has to permit one or several users, generally not experts, to solve problems regarding human expertise. Current KBS development methods help designers to define systems which function correctly from an expert competence point of view, but which have important deficiencies from an interaction with users point of view. This paper presents the basis of a methodology which views at once a co-operative approach between expert, user, computer engineer, human factor specialist and knowledge engineer for developing acceptable, useful and usable systems.

Keywords: Human-centered design, Knowledge engineering, Interactive approaches, Co-operation methodology, Human factors.

1. INTRODUCTION

At present, the problem for users and Knowledge Based Systems (KBS) interaction comes up in a crucial way since, among the available systems, few are actually used. Indeed, in most applications, KBS are only developed from expert knowledge, and end users are only considered during system integration within its environment (Mc Graw 1993).

In the first part, knowledge acquisition will be presented as a model-building approach including both elicitation and analysis. Considering this approach, it may be pointed out that expert reasoning acquisition remains one of the prominent aims of knowledge engineers. However, our attention will be devoted most particularly to the participation of KBS users throughout the developing process. Actually, it has been demonstrated by many authors that by taking human factors into account and by setting a participation approach, KBS integration in the Man-Machine system is thus made easier, thereby improving most particularly its operative ergonomics. Therefore, in the second part, a system development cooperative methodology is proposed. The third part of this paper presents an industrial example about the use of this methodology.

2. KBS METHODOLOGIES

2.1. Knowledge acquisition

The outcome of a Knowledge Based System (KBS) development project depends most of all upon the initial knowledge acquisition phase (Gaines and Shaw, 1993). As a matter in fact many studies have been aimed at improving knowledge acquisition process. Thus, knowledge acquisition no longer tries to be passed on directly from expert knowledge right to an implemented representation, but rather to formulate an intermediate expertise model, followed later on by its implementation (Johnson, 1988; Wielinga, *et al.*, 1993). While knowledge elicitation techniques are now reported to be under control, the expertise modeling phase is still deeply complex (Krivine and David, 1991).

Human sciences contribution is important to lead to a project which is aimed at modeling and broadcasting expert knowledge. Cognitive ergonomics which deals with the study of interactions between the human being and his environment, constitutes a preference inspiration source for designers. This induces to present cognitive ergonomics intervention within KBS development.

2.2. Cognitive ergonomics intervention

Cognitive ergonomics intervention could be examined on different levels as displayed by the examples hereafter, according to acceptability, utility and usability criteria (Nielsen, 1993):

- **knowledge elicitation**: the human factors specialist makes use of theoretical knowledge relative to mental processes as well as tested methods to report on the knowledge elicitation phase mainly resulting from work analysis methods (i.e. task and activity-based analysis). As regards this topic, currently active studies in cognitive sciences, but also in software and human engineering about human tasks analysis has been brought to the fore. Some graphical methods among the most famous and most used nowadays which are aimed at modeling them could be quoted: GOMS (Card, *et al.*, 1983) and its recent variations (John, *et al.*, 1994), SADT/Petri Network (Abed, *et al.*, 1991). Such methods are very useful in order to specify KBS which is actually centered on user's tasks.
- **ergonomic specification and evaluation of Man-Machine interfaces**: during the interaction determination phase between KBS and its users, the human factor specialist is an essential contribution to their specification as well as to their evaluation. To illustrate this point, in the field of ergonomics literature, hundreds of recommendations are made, aimed at the improvement of Man-Machine interfaces (Smith and Mosier, 1986; Schneiderman, 1987; Brown, 1988; Scapin, 1990). It may also be noticed that it is essential for ergonomic interface development to use a KBS potential user's model. Indeed, even though expert knowledge constitutes the KBS basis, it is nonetheless necessary to consider the subsequent use which will be made. This model will be found again in the methodology which is proposed below. Concerning KBS evaluation, the wealth of available methods should be stressed. Classifications of interactive systems evaluation methods might be seen (Wilson and Corlett, 1990; Senach 1990; Sweeney, *et al.*, 1993; Grislin, *et al.*, 1993).
- **following on-site installation**: the human factors specialist should analyze the possible constraints resulting from KBS (i.e.: work organization and task divisions) to quote but a few of them and thereby should propose adequate improvements. The previously quoted task-modeling methods will have a great importance since they will permit to compare a theoretical task model (which is defined during prior analysis) with a real model deriving from KBS user's activities. This comparison allows to enhance problems concerning its usefulness and usability (Abed, *et al.*, 1991; Millot and Roussillon, 1991).

Thus, this set of observations induces us to bring to light the dependency between the knowledge engineer and the human factor specialist. Indeed, the knowledge engineer together with the human factor specialist is to analyze the user's needs while taking the user's abilities and knowledge as well as the final system facilities and limits into account.

3. INTERACTIVE METHODOLOGY FOR KBS DEVELOPMENT

The proposed methodology is aimed at giving some answer elements to the sensitive problem concerning KBS integration within its operational environment while paying a great attention to put on to the end user's expectations.

This methodology covering the whole KBS development life-cycle describes the necessary activities to define:

- **the expertise model** to acquire and characterize expert knowledge.
- **the used model** to take the user's needs into account in order to specify Man-Machine interfaces.
- **the organization model** to take the final system operational environment into account.

3.1. Tridimensional life cycle

These three activities are integrated in to a life cycle which organizes the different KBS development phases, as shown in figure 1. This development is composed of the five following phases:

- *Problem analysis* which leads to a list of requirements. This document , which is derived from the client's requirements, points to the problem to be solved and to expectations in terms of results too.
- *Specification* which elaborates a detailed conceptual solution independently of any development means, based on the book of requirements. Specification is split up in three analysis facets: **expert knowledge analysis**, **physical environment analysis**, and **used forms specification**. The expertise model itself is completed by two models characteristic of system cooperation form with the user: the organization model and the used model.
- The *Design* phase associates these three models with a view to describing future system functionalities, i.e. to identify the main knowledge modules and their inter-relations but also define the user's interfaces.
- *Implementation* defines possible solutions from the organization and the real characteristics of software and hardware. Modules are gathered to build the final system.
- *Exploitation* confronts system and maintaining operations which integrate debugging and concern KBS evolution too.

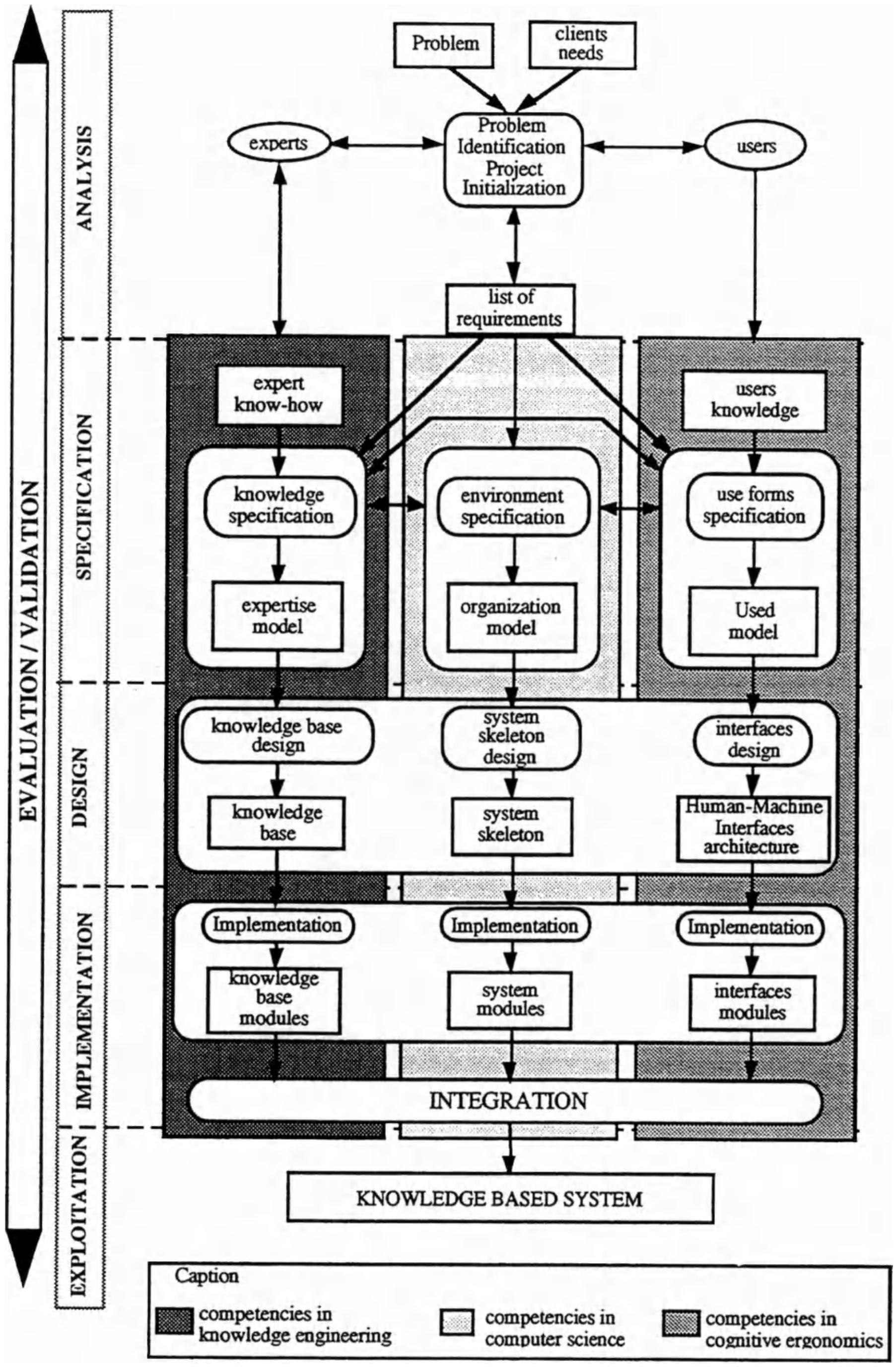


Fig. 1. Tridimensional life cycle

Validation is made all along the life cycle and constitutes by itself a current research topic (participating process, prototyping, analytical evaluation methods,...). Cf CACM (1993).

As shown in figure 1, KBS development requires three types of competencies marked in cycle by columns with different gray hues:

- Competencies in knowledge engineering to elicit and analyze expert knowledge in order to build a knowledge base.
- Competencies in cognitive ergonomics to take a future used system into account.
- Competencies in computer science to allow KBS kernel elaboration and system integration within its physical environment.

3.2. Actors involvement in life cycle

KBS development now requires the participation of actors with varied competencies and interests instead of two main interveners like before (expert and knowledge engineers). More particularly, experts and knowledge engineers, users and human factors specialists as well as computer engineers are concerned in turn during the different life cycle phases (figure 2).

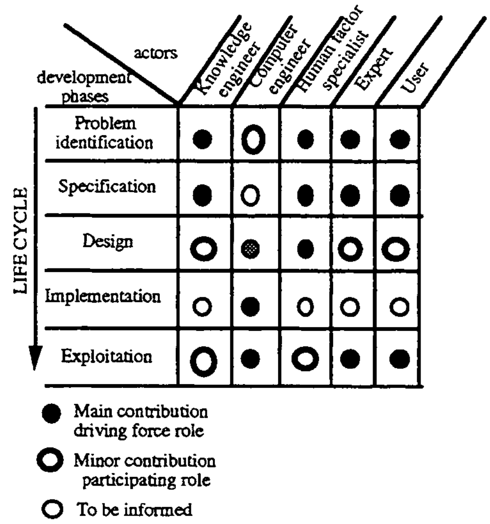


Fig. 2. Actors and life cycle phases

In KBS development, three types of specialists may be distinguished:

- The knowledge engineer who competencies are close to human sciences (psychology more particularly) elicits and analyses expert knowledge, and elaborates models.
- The computer engineer, who masters computer and artificial intelligence techniques, defines trends for physical development and concretizes system by implementing different modules.
- The human factor specialist who analyses and/or anticipates user-system interactions forms, analyses the user's needs, follows their evolution and manages adaptations until the exploitation of the system.

In order to evaluate this tridimensional life cycle, this methodological framework was tested on a real application.

4. INDUSTRIAL APPLICATION

During its activities, the L.A.M.I.H., in collaboration with the SOLLAC steel works in Dunkirk and the IRSID (Institut de la Recherche et de la Sidérurgie Française), developed a knowledge based system dedicated to help with the diagnosis of defect causes on steel rails (blooms) and plates made in the steelworks (Benkirane, *et al.*, 1991).

The setting up of continuous casting of steels produces blooms, about ten meters long, about one large and half a meter thick, from melted steel (figure 3). These blooms are later re-used through a rolling mill with the aim of transformating into plate.

This type of installation can be the cause of quality problems. Indeed, according to bloom format, melting steel quality, process supervision parameters, the users are sometimes faced with an epidemic occurrence of surface aspect defects or/and inside the blooms. The latter, observable on blooms or revealed

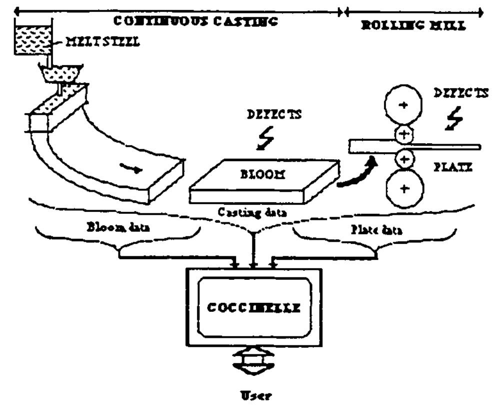


fig. 3: Bloom and plate fabrication process

on plates after lamination, impair the productivity (repairs on blooms and/or plates, scraps, manufacture stimulation ...) and come from causes sometimes difficult to identify.

The main difficulties met to make a diagnosis result from the gigantism of the installations and from the wide range of semi-finished products. These various expert's factors in front of a very large quantity of information, widely spread in time and space. It is in this context that the LAMIH, with the help of experts in continuous casting from the SOLLAC steel works in Dunkirk and the IRSID, designed and devised the "COCCINELLE" system.

KBS development follows the sequence of phases shown on figure 1. This article focuses on the specification phase with the expertise model elaboration, the used model and the organization model.

4.1. Expertise Model

Expertise model framework is based on KADS model (Wielinga, *et al.*, 1993) which considers some differentiation between knowledge types according to three levels:

- *Domain level* contains domain elements which are concerned by the problem.
- *Inference level* describes the functional point of view about domain knowledge.
- *Task level* specifies goals and the way to organize and control reasoning steps to reach these goals.

The obtained model simultaneously contains a problem solving method specification as well as a domain knowledge organization, which depends on their functions in the solving process.

In the COCCINELLE application, expert reasoning is performed through three steps which are enhance in the inference structure presented in figure 4:

- The first one identifies the population at the origin of the crisis (component) from the noted defect (problem)

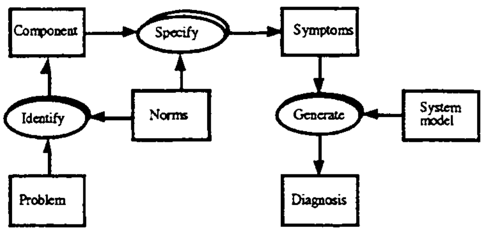


Fig. 4. "Diagnostic cause" inference

- The second one specifies the crisis characteristic clues (symptoms)
- The third one generates a list of probable defect causes (diagnostic).

4.2. Organization Model

The organization model specifies the system function, how it is split up into tasks and sub-tasks by experts. The SADT method (Structured Analysis and Design Technique) (Ross 1977) is an important contribution for elaborating this model. On the one hand it represents system functional hierarchies and data flows. On the other hand it simplifies communications between the expert, the knowledge engineer and the computer engineer.

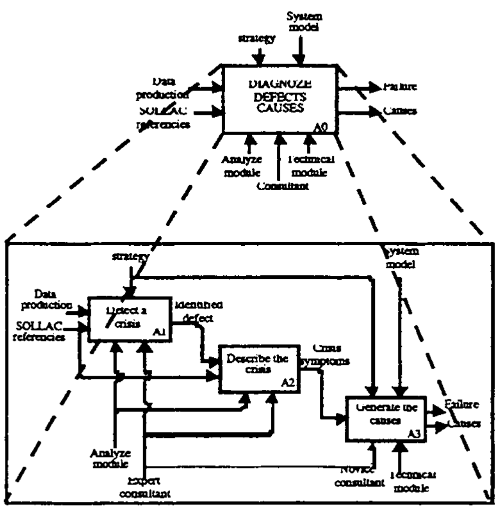


Fig. 5. Actigram of the first two application levels

The actigram in figure 5 describes a global system function and its decomposition into three sub-tasks. It permits to:

- Distinguish different modules which split up the system.
- Clearly visualize tasks management .
- Point out the user's type for the main system tasks. In the example, in order to realize detection and crisis description tasks, the consultant has to be an expert. Effectively, expertise lies in the choice of significant rates to characterize the crisis.
- Present technical constraints of system integration within its environment such as the necessary link between system and site production data and the requirement to dispose of reference thresholds which are characteristic of firm know-how.

4.3. Use model

Expert knowledge makes up KBS base but does not generally take use into account. A cognitive approach, similar to expert knowledge elicitation, allows to consider users' needs.

The cognitive ergonomics role is to disclose knowledge about users' needs and describe knowledge domain and expertise level from the user's point of view. User activity analysis takes human factors into account in order to characterize working habits, users' needs and technical constraints. The simulated expert's and user's problem-solving, for instance with the "wizard of Oz" technique (Salembier 1991), generates dialogues which are recorded and then analyzed. The knowledge engineer, in collaboration with the expert, is thus able to propose situations bringing problem types into play. Then, the human factor specialist studies interactions and assistance types.

Scenario-simulation has the advantage to elicit information for both user and expert. From the user's point of view, simulation permits to study needs in terms of information type and nature and to specify his competence and understanding levels in a given domain. From an expert's point of view, simulation enhances the aid type which is introduced by the expert in problem solving task sequences according to the simulated problems and the explanation levels the expert brings either to guide solving methods, to give precision about solving or to clarify misunderstanding points. This simulation gives occasion for the knowledge engineer to evaluate knowledge level differences between expert and user.

Elaborate used model consists in rationalizing and organizing the obtained results during simulation to describe the way by which a end-user has to represent himself system functioning. To be consistent with expertise model, used model has to keep the same organization, i.e. KADS formalism:

- the *domain level* puts in relation the user's domain knowledge with the expert's (user's own terminology, synonymous, data presentation symbolic such as textual comments, graphical objects, ...).
- The *Inference level* describes the modifications to be brought to expertise model inference level in order to take the user's competencies and goals into account.
- The *Task level* takes the user's competencies into account and navigates within the previous level. Each task is described with its goal, its control and presentation constraints. The explanations about decision-making and obtained results after the solving phase are conceivable at this level.

5. CONCLUSION

In synthesis, a KBS development methodology is proposed here, not involving expert reasoning but also:

- Taking the final system operational environment into account.
- Setting up a coherent dialogue with the user by using particularly suitable vocabulary.
- Generating explanations suited to different types of users.

Thus, KBS development should take knowledge engineering into account as regards knowledge acquisition as well as cognitive ergonomics as for the final system adaptation to the user. In both disciplines, the former conditions KBS quality while the latter is essential for the acceptance of the system by the user.

COCCINELLE is now used in industrial environment. This shows the system responds to the user's expectations and that our co-operative methodology which is the base of COCCINELLE development is validated.

REFERENCES

Abed, M., J.M. Bernard and J.C. Angue (1991). *Task analysis and modelisation by using SADT and PETRI network.* Tenth European Annual Conference on Human Decision Making and Manual Control. Liege, Belgium. 11-13 November.

Benkirane, M., M. Duribreux, and B. Houriez (1991). *Methodological knowledge elicitation approach: application to the design and realization of a knowledge based system dedicated to the technical diagnosis.* 3 rd European Conference on Cognitive Science Approaches to Process Control. Cardiff, 2-6 September.

Brown C.M. (1988). *Human-computer interface design guidelines.* Norwood, NJ: Ablex.

Card, S.K., T.P. Moran and A. Newell (1983). *The psychology of human-computer interaction.* Hillsdale, NJ: Erlbaum.

CACM (1993). Special issue about Participatory Design. *Communication of the ACM*, **36** (4), june.

Gaines B. and M.L.G. Shaw.(1993). Eliciting knowledge and transferring it effectively to a knowledge-based system. *IEEE Transactions on knowledge and data engineering.* **5** (1) p4-14, February.

Grislin M., C. Kolski and J.C. Angue (1993). *Towards an Organization of Man-Machine Interface Evaluation Techniques using a Usability Criteria Grid.* XII European Annual Conference on Human Decision Making and Manual Control, University of Kassel, Kassel, Germany. June 22-24.

John, B.E., A.H. Vera and A. Newell. (1994). Towards real-time GOMS: a model of expert behaviour in a highly interactive task. *Behaviour and Information Technology*, **13** (4), pp. 255-267.

Johnson, N.E. (1988). *Knowledge elicitation for second generation expert systems.* Proceedings of the 2nd european workshop on knowledge acquisition for knowledge based systems.

Krivine, J.P. and J.M. David (1991). L'acquisition des connaissances vue comme un processus de modélisation: méthodes et outils. In: *Intellectica,* **12**, pp. 101-137.

Mc Graw K.L. (1993). *User interfaces for knowledge-based systems. Design and evaluation.* Ellis Horwood.

Millot P. and Roussillon E. (1991). *Man-Machine Cooperation in Telerobotics: Problematics and Methodologies.* Second France Israël Symposium on Robotics, Institut National des Sciences et Techniques Nucléaires, Gif sur Yvette, April.

Nielsen, J. (1993). *Usability engineering.* Academic Press.Ross D.T. (1977). Editorial-Reflections on requirements. In: *IEEE Transactions on software engineering.* , **1**, pp. 2-5.

Ross, D.T. (1977). Editorial- Reflections on requirements. *IEEE Transactions on software engineering.* , **1**, pp. 2-5.

Salembier P., (1991) Le magicien d'Oz: une technique de simulation et d'investigation des situations d'interaction avec un système "intelligent". In: *Designing for everyone* (Y. Quéinnec & F. Daniellou, Eds), vol. 1, Paris.

Scapin D. (1990). Organizing human factors knowledge for the evaluation and design of interfaces. *International Journal of Human-Computer Interaction*, **2** (3), pp. 203-229.

Schneiderman B. (1987). *Designing the user interface: strategies for effective human-computer interaction.* Reading, MA: Addison-Wesley.

Senach B., (1990). *Evaluation de l'ergonomie des interfaces homme-machine: du prototypage aux systèmes experts.* Actes du Congrès ERGO-IA '90: Ergonomie et Informatique avancée, Biarritz, Septembre.

Smith S.L. and J.N. Mosier (1986). *Guidelines for designing user interface software.* Report EDS-TR-86-278, The MITRE Corporation, Bedford, MA.

Sweeney, M., M. Maguire and B. Shackel (1993). Evaluating user-computer interaction: a framework. *Int. J. Man-Machine Studies* , **38**, pp. 689-711.

Wielinga B.J., W. Van De Velde, G. Schreiber and S.H. Akkerman (1993). Towards a unification of knowledge modeling approaches. In: *Second generation expert systems*, (David J-M., Krivine J-P., Simmons R. Eds), pp.299-335, Spronger-Verlag.

Wilson J.R. and Corlett E.N. (eds.) (1990). *Evaluation of human works: a practical ergonomics methodology.* Taylor & Francis, pp. 890.

HUMAN-COMPUTER INTERFACE EVALUATION IN INDUSTRIAL COMPLEX SYSTEMS: A REVIEW OF USABLE TECHNIQUES

Martial GRISLIN, Christophe KOLSKI and Jean-Claude ANGUE

LAMIH, URA CNRS 1775, University of Valenciennes
BP 311 - 59304 Valenciennes, FRANCE
e-mail: grislin@univ-valenciennes.fr

Abstract: In industrial complex systems, a human-computer interface can be composed with hundreds of graphic screens and thousand variables. Human tasks are complex; HCI evaluation in this application field is difficult. During many interface developments, evaluation is often neglected by lack of knowledge about the existing techniques. Several techniques exist and may be used. The paper is focused on expert-based and analytical evaluation techniques usable in this field.

Keywords: Human-Machine Interface, Evaluation, Industrial complex systems.

1. INTRODUCTION

Contemporary literature mentions many techniques and tools for the evaluation phase —a crucial problem for the development team. Indeed many human-computer interface evaluation techniques are available today, and it is difficult to choose those really suitable to a particular situation. Moreover, their use is a long and difficult task requiring expertise and patience. In this context, our objective consists in studying and organizing the techniques that can be used for supervisory control systems with several human, ecological, economical and/or technical constraints. We are interested in graphical human-machine interface used in a control room to resolve complex problems in a dynamic situation. In the first part, this paper describes briefly the evaluation principle and properties, and the usability target data to be collected. In the second part, on the basis of a review of evaluation techniques, a set of usable human-machine interface evaluation techniques in industrial complex systems are identified and commented. This paper is focused on expert-based and analytical approaches.

2. EVALUATION PRINCIPLE

Usability evaluation aims at assessing the operator's ability to carry out his task. Senach (1990) explains that every evaluation consists in comparing an

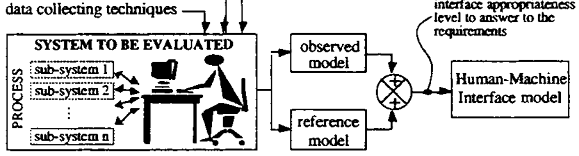


Fig. 1. Evaluation principle

observed model of the system to be evaluated with a 'reference model' (fig. 1). In interactive software evaluation, the reference model can be based on usability standards or guidelines. During the development of HCI for the supervisory systems this evaluation based on a comparison is not easy. The two models (the observed one and the reference) provide a global model of the human-machine interaction. The *observed model* comes from design methods or from empirical evaluation. The *reference model* comes from task analysis methods, i.e., human oriented design. This evaluation principle helps the development team and enables to take into account the usability guidelines during the cycle. The realization and the validation of used models are the main difficulty in providing good evaluation. These descriptions are formal models of the dialogue between the human supervisor and the machine. Another important difficulty concerns the availability model: formalizing the man-machine interaction is not easy and there is often a gap between theory and reality. The problem is to know whether one can trust an evaluation based on a representation that may not reflect reality. The choice of a HMI model for evaluation is determined by the context. Thus the selection of adequate properties to evaluate quality depends not only on the evaluation objectives but also on the operators' characteristics and on the task requirements. The range of functions and the user-friendliness of an interface are significant properties, defined in the terms *utility* and *usability*. *Utility* determines whether the interface enables the user to achieve his goals, in terms of functional abilities, performance and aid quality. *Usability* describes the quality of the interaction between human and machine (ease of learning and use) and reflects the consistency of the design.

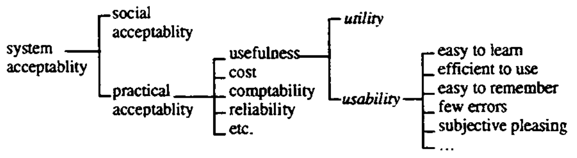


Fig. 2. Principal evaluation data (Nielsen 1993)

An analysis of the interface and the interaction brings to light a set of target variables, i.e., basic data to be collected. There are two families of principal variables: a social and a practical acceptability (fig. 2). This second class includes current production, cost, reliability constraints, and usefulness of interface. *Utility* and more particularly *usability* is located in this criterion. Contemporary literature offers numerous usability and ergonomic recommendations: guidelines or standards. They can be used during evaluation as *usability* criteria to observe or measure. Relying on several works, Bastien and Scapin (1993) discussed a set of criteria and sub-criteria that take into account such recommendations (table 1). The whole criteria help the evaluation team to estimate the interface usability quality and to make the necessary decision concerning modification and/or improvement.

Table 1 Evaluation usability criteria, basic criteria in bold and numbered (Bastien and Scapin 1993)

Criteria	*Principle and objective*
Guidance C1 C2 C3 C4 C5	Guidance facilitates discovery and usage by helping the user to identify further states and actions possible. **Prompting (C1)**: provide guidance about what the user can do. **Grouping of items by localization (C2)**: relationships between displayed information (organizing items into hierarchic lists, distinct areas, etc.), and **Grouping of items by format (C3)**: distinguishing between functional areas, distinctive label of categories, color coding, etc. **Immediate feedback (C4)**: results of users' action. **Clarity (C5)**: facilitate legibility
User workload C6 C7 C8	Workload concerns all interface elements that play a role in the reduction of the users' perceptual or cognitive load, and in the increase of the dialogue efficiency. This criterion includes two first areas. **Concision (C6)**: individual items (data, commands, etc.) use the shortest possible items to be entered by the user (short items); and **Minimal actions (C7)**: set of items (procedures) for which it is suggested to allow only the shortest, most necessary action. The third area is the **Mental load (C8)** or information density. It concerns the users' workload from a perceptual and cognitive point of view regarding the whole set of information presented.
Explicit control C9 C10	Control in user entries favors focusing on the user's actions and therefore limits ambiguities and errors. Two sub criteria are defined. **Explicit actions (C9)**: unambiguous user specification of an action or set of action; and **User control (C10)**: allow processing of only what the user is asking for.
Adaptability C11 C12	The adaptability of the system refers to its capability to behave contextual and according to the users' needs and preferences: provide various ways for the user to interact. One general aim is **Flexibility (C11)**: the capability of the interface to adapt to various user actions. One more specific aim is flexibility applied to users' experience, that is, **Users' experience management (C12)**.
Error management C13 C14 C15	The aim is to avoid errors as much as possible. When errors occur, identify the error(s), its location, and means to correct the error(s). **Error protection (C13)**: means available to detect and prevent data entry errors, command errors, or actions with destructive consequences. **Quality of error messages (C14)**: phrasing and content of error message, their relevance, readability, and specificity about the nature of the errors (syntax, format) and the actions needed to correct them. **Error correction (C15)**: means available to the users to correct their errors.
Consistency (C16)	Consistency refers to the way interface design choice (codes, naming, formats, procedures, etc.) are maintained in similar contexts, and are different when applied to different contexts. It concerns location (similar window location, message area location), format (similar screen formats), syntax, procedures (similar order in command language syntax, and procedures for menu selection), naming (similar naming of commands or menu option, abbreviations, and labels).
Significance of codes (C17)	It concerns the association between an individual term and/or sign (label, command, etc.) with its reference. Codes and names are significant to the users when there is a strong semantic relationship between such codes and the items or actions they refer to. Not the coding related to the distinction or grouping of several items belonging to the same class.
Compatibility (C18)	Compatibility refers to the match between users' characteristics (memory, perception, customs, skills, age, expectations, etc.) and task characteristic at the one hand, and the organization of the output, input, and dialogue for given application, on the other hand. It's also concerns the coherence between environments and between applications.

3. REVIEW OF USABLE EVALUATION TECHNIQUES

HCI evaluation depends on the characteristics of the situation to be evaluated. When the interface to be evaluated exists, user behavior data analysis methods are recommended. Data are collected with observations or measures on real site or on laboratory. Those methods consist in realizing usability tests with real users or selected subjects for experiment. This evaluation type is realized *a posteriori* on an already used interface or a prototype. It provide opinion about their use. In this first case, several methods and techniques exist. They can be classified into three approaches (fig. 3): (1) empirical approaches, (2) observing approaches, (3) expert-based approaches. When the interface to be evaluated does not exist, analytical methods can be used. They allow an *a priori* analysis, i.e. provide an interface valuation before his real use. In this second case, the corresponding techniques can be classed into two approaches (fig. 3): (3) some expert-based approaches (already cited above), (4) analytical approaches. Analytical evaluation allows to describe an interface that responds to the operator requirement in terms of utility and usability. It takes into account the human oriented design recommendation and can be supported by an appropriate architecture of human-machine dialogue. Those methods can be used during interface specification and design, to check *a priori* the proposed interface usability quality. The usable techniques or tools for human-computer interface evaluation come from various fields, such as software engineering, knowledge elicitation, ergonomics studies or psychology. Many methods exist. This classification synthesizes some evaluation studies (Maguire and Sweeney, 1989; Senach, 1990; Whitefield *et al.*, 1991; Balbo and Coutaz 1992), with an update. We take into account only those usable for complex systems. We used those techniques currently

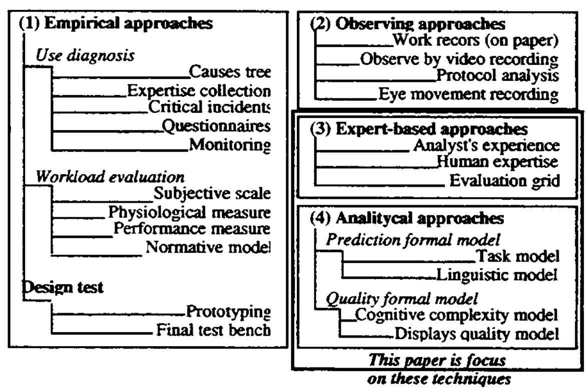


Fig. 3. HCI evaluation techniques classification

in real and/or simulated industrial situation (chemical process, nuclear, automobile and rail transport, submarine...). This paper is focused on our last research in evaluation: an expert-based and an analytical approach. The next sub-section gives a short discussion about the techniques of other approaches (for more details, see Kolski, 1993).

3.1 *Empirical and observing approaches*

Empirical approaches consist in measuring representative users' performance with experimental methods and tools. The information from *Use diagnosis* technique facilitates design and improvement decision making. But, they do not take into account the temporal evolution of the system and the identification of the failure sequence is difficult for complex interactive systems. The *Workload evaluation* is based on the user's cognitive work observation. This qualitative or quantitative measuring of the user's activity contributes to the usability evaluation in line of the interactive system. Several methods and techniques are proposed (Wierwille and Eggeneier, 1993). The analysis of those methods leads to the conclusion that workload evaluation may be based on the use of several indicators: (i) a description of the user's work gives information about their different strategies, (ii) a subjective scale give a reference evaluation of the mental workload, and (iii) a physiological indicator informs of the user's capability and his vigilance level. Several recent works, in the LAMIH, devote an important development in a temporal model based method (Berger *et al.* 1989; Riera and Grislin 1995), and an eye scanning movement measure (Simon *et al.* 1993). The *Design test* methods realize an evaluation with an iterative cycle during the whole design process. At every development step, the interface is tested which generate modification, then tested in a new version, and so on until it is satisfactory. This evaluation engineering is very interesting for process control. It permits to take automatically into account the criteria and constraint related to the process control (production, economy...). But, those methods imply the use of a powerful and realist simulator, that is very expensive. *Observing approaches* group methods only based on the observation of specific aspects concerning the HCI. It consists in collecting representative data of the interaction, and to analyze those user activity traces. The collected data describe the event occurring during the user's activity: failures, orders modifications, stopping, takeoff, errors and correction actions, etc. In the supervisory of an industrial complex system with graphical aids tools, the operator uses mainly his visual sense. In this case, analysis of eye scanning movement is logically integrated in an evaluation process. This technique is usable in parallel of other analysis tools. Eye movement record is used in the LAMIH for the task analysis in the interface specification and/or improvement (Abed, 90).

Table 2 Expert-based approaches

Techniques	*References*	*Usability criteria*
Analyst's experience	Sperandio, 1991	
Human expertise	Molich *et al.* 90	every criteria
Evaluation grid	Smith and Mosier, 1986; Ravden and Johnson, 1989	every criteria

3.2 *Expert-based approaches*

More often, the HCI evaluation requires an expert analysis. Expert-based approaches consist in comparing the interface with a human-computer interaction informal model (table 2). The **Analyst's experience** consists for the analyst in replacing the operator in his job. This technique let us see the difficulties met by the user. It is recommended for a simple task (office automation), but can rarely used when the task is complex. In a control room, this method raises a huge problem: it is not possible for the evaluator to use the supervisory system in line. The type of evaluation more currently used is the **Human expertise**. It is efficient and simple to organize for a rapid diagnosis of design errors. Every criterion is concerned as human-factor experts have to be able to use them. But this type of method is based on an individual reference model acquired with experience. Every specialist is focused on specific characteristic of the interface and bases his evaluation on his personal model. To get a more exhaustive approach of the HCI problems and to avoid a focusing error, it is necessary to take several different specialists, and therefore it implies an important cost. the **Evaluation grid** is very interesting for the measurement of the interface usability. This technique has to take into account every characteristic of the interface (an exhaustive list of criteria). It obtains an excellent result for a comparative evaluation of interfaces and prototypes. A, evaluation grid is usable by designers, human factor specialists, and sometimes by the final users. But they are difficult to understand for non-expert of human factors: recommendations are subjective and/or can be interpreted differently, and the notation is difficult to analyze. The grid is more adapted for a human-factor expert because they have already a mental model of the user's work.

3.3 *Analytical approaches*

Analytical approaches consist in realizing the evaluation with a model of the interface or of the human-computer interaction. So, it is necessary to identify the pertinent variable, according to the evaluation objective and the task requirement (utility and usability). Then, the collected data analysis require measurement scale to be used during the comparison (a, exploitable and objective notation). Results may be integrated into a global appreciation of HCI. Those approaches use formal models and objective metrics. The models are used to predict some aspects of the interaction: task hierarchy, developing of the user action, time... The metrics allow to measure with objectivity some quality aspect of the interaction: consistency, compatibility with the mental scheme of the user, display quality... Those abstract representations allow to predict the performance of the system, which cannot be obtained with empirical approaches.

Predictive formal model: The human task modeling is particularly important if it facilitates the communication between the different agents of the development team: showing the user's assistance requirements, choosing the different aid's function... The modeling can be realized with different type of description. That model is based on a goal, sub-goal description (fig. 4). Table 3 lists the most known ones among those which are usable for interactive applications.

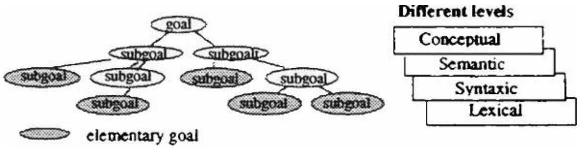


Fig. 4. Task description as a goal sub-goal hierarchic

Task model. The task models consist in a hierarchical decomposition of task in order to formalize the cognitive activity of the user. In this type of model, the task is decomposed through a refining process starting from global task or from elementary task (descending and ascending analysis). The analysis provides a formal support for the predictive evaluation of the performance (a measurable description of the user's behavior). **HTA** is a hierarchical model of the user's task. Every task corresponding to a planning, a sequence (or a procedure), can be described for example with a petri net and execution conditions. **GOMS** predicts the length of the task realization, and describes the choice of methods. With its formal approach, GOMS has opened some interesting perspectives and has lead to defining of other models (e.g., KLM). The first version of that model was validated for very simple tasks (office automation). But this model is limited because it takes into account neither the user's errors, nor the resolution of the problem. This limitation does not allow its direct using for the prediction of the user's efficiency when he realizes a complex task. But we can note that a recent validation of highly interactive behaviors has been realized with an extended model tell CPM-GOMS based on a critical path method (Gray *et al.* 1993). The INRIA laboratory has developed the **MAD** method to facilitate the user's task description. The obtained task model shows the difference between functional logic and using logic. The user's characteristic is taken into account rather than the application functioning logic and than the computer constraint. This approach is very interesting, but there are some limitations. The interruption of the activity is not visible in the description The human errors and/or procedures are not taken into account in the description. Note that this formalism may be improved in the new version of MAD. Therefore some research in the LAMIH has developed a method based on **SADT and Petri networks**. In this method, SADT formalism realizes the functioning decomposition with a top down and a hierarchical point of view. Every module was structured according to the SADT diagram decomposition. In the end of

Table 3 Example of predictive formal model: task and linguistic model

Techniques	*References*	*Usability criteria*
HTA	Annet *et al.*, 1971; Stammer *et al.* 1990	C1 C7 C8 C11 C12
GOMS KLM	Card *et al.*, 1983; Irving *et al.*, 1994	C1 C7 C8 C11 C12
MAD	Scapin and Pierret-Golbreich, 1989	C1 C7 C8 C11 C12
SADT-Petri or TOOD	Abed, 1990; Mahfoudi *et al.*, 1995	SADT: C1 C7 C8 C11 C12 Petri: C4 C9 C10 C13
ALG	Reisner, 1984	C1 C7 C8 C11 C12
CLG	Moran, 1981	C1 C4 C8 C9 C10 C12 C13 C15 C16 C18
TAG et ETAG	Payne and Green, 1986; Tauber, 1990	C1 C7 C8 C11 C12

these functioning decomposition phases, the decision behavior is described using synchronized petri network. The obtained model is next used during the usability evaluation for the user's activity analysis. This method has recently evolved with TOOD: Task Object Oriented Description (Mahfoudhi *et al.* 1995).

Linguistic model: The linguistic model can describe the interactive task between the user and his interface, in order to evaluate the command language and its consistency. Those models explain the human-machine interaction structure with a grammar. For instance, **ALG** proposes to describe the interface procedures used with production rules. But with the hierarchical description only, this grammar neither shows the relationship between the display, nor the dynamic of dialogue. **CLG** is a design oriented approach, usable also for the evaluation. The task description is transformed into a set of models realized on six levels: task and semantic for goals and sub-goals, syntax and interaction for the communication between the user and the system, presentation and input-output unit for physical component. This formalism allows to locate where the evaluation decision was critical for the user. For example: the interface efficiency is measured with an estimation of the action speed or the apprenticeship duration. **TAG** model describes the mental representation of the interaction language. The task is broken up into simple routine sub-tasks (or action) which the user can execute without decision making. But TAG is limited to simple tasks, i.e., the description of specific system function. There is no relationship between the different simple tasks into the upper levels.

Conclusion. This hierarchical description of the task in elementary sub-tasks shows the abilities and the information necessary to the user. Tasks are structurally split up to describe the parallelism and the sequential ordering of the procedure. Today, this type of model is in evolution and can be used for design test in the place of empirical studies with prototype. But they limit the evaluation because they predict only the user's performance corresponding to the defined or analyzed model. For a simple task, the linguistic models give a good result. But after our first tests for interface evaluation in a complex industrial system control, those models don't give any more information than task model. The hierarchical aspect of the interaction is to be keep. The description of the human-machine dialogue as a software language (interesting for automation) help to detect some problem: consistency, interaction difficulty, compatibility of the apprenticeship. But some research might be developed to make them valid.

Quality formal model: The quality formal models concern the measurement properties of the interface in a more practical view than the above model. They are currently based on a usability and ergonomic knowledge base. Table 4 lists some of them, non-exhaustively. The display evaluation of interactive system is based on the principle that some problems occurring during the use can be avoided *a priori*.
The first type of automatic evaluation systems can analyze static information presentation. The **Display analysis program** measures the quality of alphanumeric language exclusively. The prototypes **SYNOP** and **KRI/AG** evaluate, respectively, the industrial mimic displays and the dialogue interfaces in a graphical environment. SYNOP modifies directly the screen, and KRI/AG provides a usability recommendation written like comments in textual reports. These two tools are limited: display assessment one to one without global view, non-fulfilment of the knowledge base. KRI/AG system, more recent, interprets the standard UIL format (User Interface Language). This characteristic is very interesting for an automatic evaluation, and can be integrated in a future evaluation tool. This first type of approach shows the possibility to automate the many design guidelines currently proposed in a daunting presentation (manual and guide on paper).
Other automatic evaluation system based on knowledge base are able to generate the interface or specification. From a description of the system, that method aims at obtaining a display respecting basic recommendation. **APT** is too limited for industrial interface design because it can only represent and test static numeric data. The Perlman program can describe the display structure in different abstraction levels. The aspect link to the presentation are formalizes as a network, with relationship between the display structural properties and presentation rules. This type of approach is too limited as a design help of the aspect based on the presentation, without analyzing the use, nor the dynamic. Today, the automatic interface generator might concern the dynamic evaluation of the HCI. **Ergo-conceptor**, developed on the LAMIH, is based on the description of the application (i.e., an industrial system to be supervised). This tool assists the designer with some proposition of representation mode, supposedly adapted to the control task. The designer can choose directly that prescription and generates interactively the different mimic displays. This type of evaluation systems are interesting because they can automate several designs and realization steps. But there are also some problem: non-fulfilment of the knowledge base, difficulty to realize a good description of the application.

Table 4 Example of quality formal model: display quality model and automatic generation

Techniques	*Reference*	*Usability criteria*
Display Analysis Program	Tullis, 1986	C2 C3 C5 C8
SYNOP	Kolski, and Millot, 1991	C2 C3 C5 C8 C16
KRI/AG	Lowgren and Nordqvist, 1990	C2 C3 C5 C8 C16
APT	Mackinlay, 1986	C2 C3 C5 C6 C8 C16 C18
Data structure	Perlman, 1987	C2 C3 C5 C6 C8 C16 C18
Ergo-conceptor	Kolski and Moussa, 1991	C2 C3 C5 C6 C8 C16 C18

Conclusion: Nowadays some researches are still carried out about some of the models which can not directly be used for an evaluation in an industrial situation. Their using demands an important and long formalism work and very little information was given about the method to realize the test during design. An other important point is that an evaluation tool of HCI can not be limited to an evaluation of exclusively every usability aspect. It might estimated the advantages and weaknesses. The analysis might concerns the input and output command language, and the screen organization. The different works presented in this section suggest that it is difficult to develop a tool able to evaluate every interface. A really satisfactory tool must be constituted on an evaluation environment.

4. CONCLUSION

Many HCI evaluation methods and techniques exist. Therefore some of them are usable for human-machine interface used in simple task (e.g., office automation). For several years, we have been testing some of those methods in industrial situations. The empirical, observing and expert-based approaches give relatively good results (of course some adjusting is still needed). The realization of a controlled experiment and the recording, and then the data analyses take a very long time. It is currently non-compatible with the time requirement and delay of the system development. And moreover, the final users are rarely available or accessible, and the designers are not already convinced of the utility of those early-tests. In fact, the evaluation is considered more as a corrective function than as a generative information source for a design solution. Concerning the analytical approaches, someone of them raise a problem because of the dynamic of the situations to be evaluated. These are the aim of our present research in complex situations.

REFERENCES

Abed M. (1990). *Contribution à la méthodologie de la tâche de conception par des outils de spécification exploitant les mouvements oculaires*. Ph. D. Thesis, Valenciennes.

Annet J., K.D. Duncan, R.B. Stammers, M.J. Gray (1971). *Task analysis* (HMSO London).

Balbo S. and J. Coutaz, (1992). *Un pas vers l'évaluation automatique des interfaces homme-machine*. ERGO IA'92, 7-9/10, Biarritz, 367-382.

Bastien J.M.C. and D.L. Scapin (1993) *Ergonomic* *criteria for the evaluation of Human-computer Interfaces*. Rapport INRIA n°156, Juin, 79p.

Berger T., B. Riera, J.C. Angue, P. Millot (1989). *Real time assessment of human operator workload placed in a simulated multitask context of a monitoring and supervising process*. ESA workshop, Human factor engineering: a task-oriented approach, 21-23/11, 13 p.

Card S., T.P. Moran, A. Newell (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Lawerence Erlbaum Associates.

Gray W.D., B.E. John, M.E. Atwood (1992). *The precis of project ernestine or an overview of validation of GOMS*. CHI'92, ACM Conference on Human factors in Computing Systems, Bauersfeld *et al.* (Ed.), Monterey, California, May 3-7, 307-312.

Irving S., P. Olson, J.E. Irving (1994). *A GOMS analysis of the advanced automated cockpit*. CHI'94 , ACM Press, Addison-Wesley.

Kolski C. (1993). *Ingénierie des interfaces homme-machine, conception et évaluation*. Hermes (Ed.), Paris, 372 p.

Kolski C. and F. Moussa (1991). Une approche d'intégration de connaissances ergonomiques dans un atelier logiciel de création d'interfaces pour le contrôle de procédé. *Quatrièmes Journées Internationales: le Génie Logiciel et ses applications*, Toulouse, 9-13 décembre.

Kolski C. and P. Millot (1991). A rule-based approach to the ergonomic "static" evaluation of man-machine graphic interface in industrial processes. *Int. J. of Man-Machine Studies*, 35, 657-674.

Lowgren J. and T. Nordqvist (1990). *A knowledge-based tool for user interface evaluation and its integration in a UIMS*. In Diaper *et al.* (Eds.), INTERACT'90, 395-400, North-Holland.

Mackinlay J. (1986). Automating the design of graphical presentation of relationnal information. *ACM Transactions on Graphics*, 5 (2), 110-141.

Maguire M. and M. Sweeney (1989). *System Monitoring: Garbage Generator or Basic for comprehensive Evaluation system?* The fifth conf. of the British Computer Society HCI, Specialist Group, University of Nottingham, september 5-8.

Mahfoudhi A., M. Abed M., J.C. Angue (1995). *A new methodology for Man-Machine Systems analysis and design Design*. Industrial Ergonomics and Safty Conference, Seattle, Washington USA, June 13-16.

Molich R., J. Nielsen (1990). Improving A Human Computer Dialogue. *Communications Of The Acm*, 33, 338-348.

Moran T.P. (1981). The Command Language Grammar: a representation of the user interface of interactive computer systems. *Int. J. of Man-Machine Studies*, 15, 3-50.

Nielsen J. (1993). *Usability Engineering*. Academic Press, Boston, 358 p.

Payne S.J., T.R.G. Green (1986). Task-action grammars: A model of the mental representation of task languages. *Human-Computer Interaction*, 2, 93-133.

Perlman G. (1987). *An axiomatic model of information presentation*. Human Factors Society Meeting, New York, 1219-1223.

Ravden S. and G. Johnson (1989). *Evaluating Usability Of Human-Computer Interface*. Chichester: Ellis Horwood.

Reisner P. (1984). Formal grammars as a tool for analysing ease of use: some fundamental concepts. In Thomas and Schneider (Eds). *Human Factors in Computer Systems*. Norwood, N.J.: Ablex.

Riera B., M. Grislin, P. Millot (1994). *Methodology To Evaluate Man-Car Interfaces*. IFAC/IFORS Transportation Systems: Theory and Application of Advanced Technology, Tianjin, China, August 24-26, 459-469.

Scapin D.L. and C. Pierret-Golbreich (1989). *MAD : Une méthode analytique des descriptions des tâches*. Colloques sur l'ingenierie des IHM, 131-148, Sophia Antipolis, 24-26 Mai.

Senach B. (1990). *Évaluation ergonomique des interfaces homme-machine : une revue de la littérature*. Rapport INRIA n°1180, 75 p.

Simon P., T. Berger, J-C. Angue (1993). *Is quantitative of eyse scanning movement relevant of eye scanning movements/*. International Industrial Ergonomics and Safty Conference'93, Copenhague, Danemark, june 8-10.

Smith S.L. and J.N. Mosier (1986). *Guideline for designing user interface software*. Report MTR-10090, ESD-TR-86-278, Bedfort, MA: The MITRE Corporation.

Sperandio J.C. (1991). Les méthodes d'analyse du travail en psychologie ergonomique. In *La recherche en Psychologie (Domaine et méthodes)* de J.P. Rossi, Dunod (Ed.).

Stammer R.B., M.S. Carey, J.A. Asteley (1990). Task analysis. In Wilson et Corlett (Eds), *Evaluation of human work , a practical ergonomics methodology*. Taylor & Francis Ltd, 135-160.

Tauber M. (1990). *ETAG: Extended Task Action Grammar-A language for the description of the user's task language*. In Diaper *et al.* (Eds), INTERACT'90, 163-168, North-Holland.

Tullis T.S. (1986). *Display analysis program (Version 4.0)*. Lawrence, Kansas, The Report Store.

Whitefield A., F. Wilson, J. Dowell (1991). A framwork for human factors evaluation. *Behaviour & Information Technology*, vol 10, n°1, january-february, 65-79.

Wierwille W.W. and F. Eggeneier (1993). Recommendation for mental workload measurement in a test and evaluation environment. *Human Factors*, 35 (2), 263-281.

USABILITY EVALUATION: AN EMPIRICAL VALIDATION OF DIFFERENT MEASURES TO QUANTIFY INTERFACE ATTRIBUTES

Matthias Rauterberg

Work and Organisational Psychology Unit
Swiss Federal Institute of Technology (ETH)
Nelkenstrasse 11, CH–8092 Zuerich
Tel: +41-1-63-27082, Email: rauterberg@rzvax.ethz.ch

Abstract: One of the main problems of standards (e.g., DIN 66234, ISO 9241) in the context of usability of software quality is, that they can not be measured in product features. We present a new approach to measure user interface quality in a quantitative way. First, we developed a concept to describe user interfaces on a granularity level, that is detailed enough to preserve important interface characteristics, and is general enough to cover most of known interface types. We distinguish between different types of 'interaction points'. With these kinds of interaction points we can describe several types of interfaces (CUI: command, menu, form-fill-in; GUI: desktop, direct manipulation, multimedia, etc.). We carried out two different comparative usability studies to validate our quantitative measures. The results of one other published comparative usability study can be predicted. Results of six different interfaces are presented and discussed.

Keywords: user interfaces, utility functions, testability, quantization.

1. INTRODUCTION

One of the main problems of standards (e.g., DIN 66234, ISO 9241) to quantify software quality of usability is, that they can not be measured in product features (Kirakowski and Corbett, 1990). Four different views on human computer interaction to measure interactive qualities currently exists (see also Bevan, et al., 1991, p. 651; Rengger, 1991).

(1) The *interaction-oriented view:* usability quality is measured in terms of how the user interacts with the product ("usability testing"). This view is the most common one. All kinds of usability testing with "real" users are subsumed in this category (IFIP, 1981).
(2) The *user-oriented view:* usability quality is measured in terms of the mental effort and attitude of the user ("questionnaires" and "interviews").
(3) The *product-oriented view:* usability quality is measured in terms of the ergonomic attributes of the product itself (quantitative measures).
(4) The *formal view:* usability is formalised and simulated in terms of mental models (formal concepts). Karat (1988) describes formal methods in the context of "theory-based" evaluation.

The interactive qualities of user interfaces currently are quantified in the context of *interaction-oriented view* and *user-oriented view*, but these both approaches are time consuming and more or less expensive (Jeffries and Desurvire, 1992).

2. A DESCRIPTIVE CONCEPT OF INTERACTION POINTS

We present a new approach to measure user interface quality in a quantitative way. First, we developed a concept to describe user interfaces on a granularity level, that is detailed enough to preserve important interface characteristics, and is general enough to cover most of known interface types (command language, CUI, GUI, multimedia, etc.). Different types of user interfaces can be quantified and distinguished by the general concept of "interaction points". Regarding to the interactive semantic of "interaction points" (IPs), different types of IPs must be discriminated (see also Denert, 1977).

An interactive system can be distinguished in a dialog and an application manager. So, we distinguish between dialog objects (DO, e.g. "window") and application objects (AO, e.g. "text document"), and dialog functions (DF, e.g. "open window") and application functions (AF, e.g. "insert section mark"). Each function f∈FS, that changes the state of an application object, is an application function. All other functions are dialog functions (e.g., window operations like move, resize, close). The complete set of all description terms is defined in Table 1.

Table 1 The interactive space (IS) consists of the object space (OS) and the function space (FS); FS can be distinguished in perceptible and hidden interactive functions (PF and HF).

$IS := OS \cup FS$	[interaction space]
$DC \in IS$	[dialog context]
$OS := PO \cup HO$	[object space]
$FS := PF \cup HF$	[function space]
$PO := PDO \cup PAO$	[(perceptible) representations of objects]
$HO := HDO \cup HAO$	[hidden objects]
$PF := PDFIP \cup PAFIP$	[(perceptible) representations of functions]
$HF := HDFIP \cup HAFIP$	[hidden functions]
$PDFIP := \{(df,pf) \in HDFIP \times PF: pf = \delta(df)\}$	[(perceptible) represented DFIP]
$PAFIP := \{(af,pf) \in HAFIP \times PF: pf = \alpha(af)\}$	[(perceptible) represented AFIP]
$IP := DFIP \cup AFIP$	[interaction points]
$DFIP := PDFIP \cup HDFIP$	[IPs of dialog functions]
$AFIP := PAFIP \cup HAFIP$	[IPs of application functions]
δ := mapping function of a df $\in$ HDFIP to an appropriate pf $\in$ PF.	
α := mapping function of an af $\in$ HAFIP to an appropriate pf $\in$ PF.	
$PDO := \{(do,po) \in HDO \times PO: po = \mu(do)\}$	[(perceptible) represented DO]
$PAO := \{(ao,po) \in HAO \times PO: po = \nu(ao)\}$	[(perceptible) represented AO]
μ := mapping function of a dialog object do $\in$ DO to an appropriate po $\in$ PO.	
ν := mapping function of an application object ao $\in$ AO to an appropriate po $\in$ PO.	

A dialog context (DC) is defined by all available objects and functions in the actual system state. If in the actual DC the set of available functions changes, then the system changes from one DC to another. All dialog objects (functions, resp.) in the actual DC are perceptible (PO, PF) or hidden (HO, HF). Four different mapping functions relate perceptible structures to hidden objects or functions.

Each interaction point (IP) is related to at least one interactive function. If both mapping function's δ and α are of the type 1:m(any), then the user interface is a command interface. If both mapping function's δ and α are of the type 1:1, then the user interface is a menu or direct manipulative interface where each f∈FS is related to a perceptible structure PF (see Figure 2). The perceptual structure (visible, audible, or tactile) of a function (PF) can be, e.g., an icon, earcon, menu option, command prompt, or other mouse sensitive areas.

The intersection of PF and PO is sometimes not empty: $PF \cap PO \neq \emptyset$. In the context of graphical interfaces icons are elements of this intersection, e.g., PDFIP "copy" ≡ PDO "clipboard", PAFIP "delete" ≡ PAO "trash". Each interaction point (IP) is related to at least one interactive function (see Figure 1).

One important difference between a menu and a direct manipulative interface can be the "interactive directness". A user interface is 100% interactively direct, if the user has fully access in the actual dialog context to all f∈FS (Laverson, et al., 1987). Good interface design is characterised by optimising the multitude of DFIPs (e.g. "flatten" the menu tree; see Paap and Roske-Hofstrand, 1988) and by allocating an appropriate PDFIP to the remaining HDFIPs.

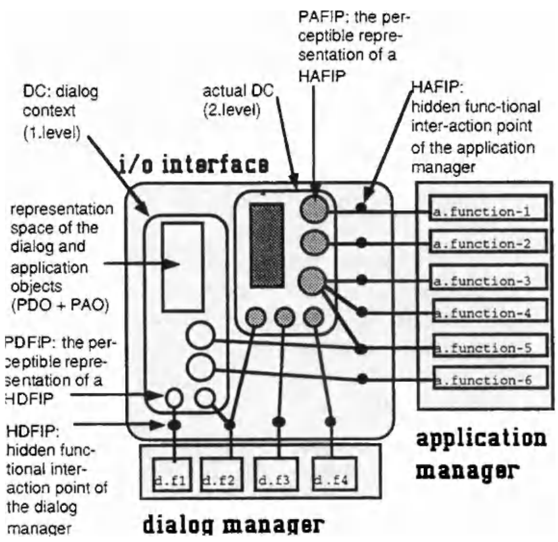


Fig. 1. A schematic presentation of the I/O interface, the dialog and the application manager of an interactive system with a menu tree of two levels.

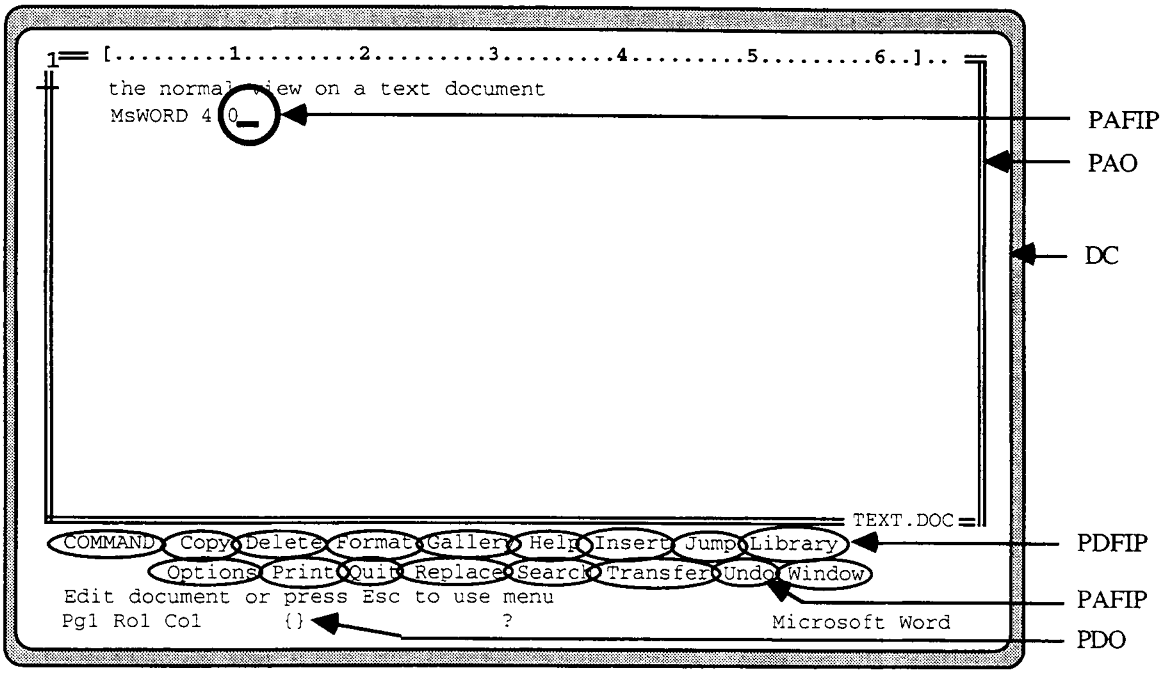


Fig. 2. An actual dialog context (DC) of the text processing program MsWord with the representation space of the interactive object (PAO: text document; PDO: clipboard), and the representation space (PF: marked by circles) of the interactive functions (PAFIP: text entry point, undo; PDFIP: menu options).

3. FOUR QUANTITATIVE MEASURES OF INTERFACE ATTRIBUTES

To estimate the amount of "feedback" of an interface a ratio is calculated: "number of PFs" (#PF = #PDFIP + #PAFIP) divided by the "number of HFs" (#HF = #HDFIP + #HAFIP) per dialog context. This ratio quantifies the average "amount of functional feedback" of the function space (FB; see Formula 1). We abbreviate the number of all different dialog contexts with D. A GUI has often a very large number of DCs. To handle this problem we take only all task related DCs into account. Doing this, our measures will give us only a lower estimation for GUIs.

The average length of all possible sequences of interactive operations (PATH) from the top level dialog context (DC, e.g., 'start context') down to DCs with the desired HAFIP or HDFIP can be used as a possible quantitative measure of "interactive directness" (ID, see Formula 2). The measure ID delivers two indices: one for HAFIPs and one for HDFIPs. A PATH has no cycles and has not more than two additional dialog operations compared with the shortest sequence. An interface with the maximum ID of 100% has only one DC with path lengths of one dialog step. We abbreviate the number of all different dialog paths with P.

Functional feedback: $$FB = 1/D \sum_{d=1}^{D} (\#PF_d \,/\, \#HF_d) \quad * 100\% \qquad (1)$$

Interactive directness: $$ID = \left\{ 1/P \sum_{p=1}^{P} \operatorname{lng}(PATH_p) \right\}^{-1} * 100\% \qquad (2)$$

Application flexibility: $$DFA = 1/D \sum_{d=1}^{D} (\#HAFIP_d) \qquad (3)$$

Dialog flexibility: $$DFD = 1/D \sum_{d=1}^{D} (\#HDFIP_d) \qquad (4)$$

To quantify the flexibility of the application manager we calculate the average number of HAFIPs per dialog context (DFA; see Formula 3). To quantify the flexibility of the dialog manager we calculate the average number of HDFIPs per dialog context (DFD; see Formula 4). A modeless dialog state has maximal flexibility (e.g., "command" interfaces, or Oberon; Wirth and Gutknecht, 1992).

Let us apply the five measures to our example in Figure 1. The average amount of functional feedback is:
$FB = (4/4 + 6/8) /2 * 100\% = 87.5\%$.
The average amount of interactive directness is:
$ID_{HAFIP} = ((2*1 + 5*2) / 7)^{-1} * 100\% = 58.3\%$;
$ID_{HDFIP} = ((2*1 + 3*2) / 5)^{-1} * 100\% = 62.5\%$.
The average amount of flexibility is:
$DFA = (2 + 5) / 2 = 3.5$ and
$DFD = (2 + 3) / 2 = 2.5$.
To interpret the results of our measures appropriately, we need empirical studies.

4. RESULTS AND DISCUSSION

We carried out two different comparative usability studies to validate our measures (Rauterberg, 1992; Brunner and Rauterberg, 1993). A third external comparative study (Grützmacher, 1988) was used for a cross validation. All three investigated software products have the same application manager, but two different dialog managers each.

4.1 *Results of experiment-I*

We compared an old CUI-interface of a relational database management system with a new GUI-interface (Rauterberg, 1992). The main result of this empirical investigation was, that the mean task solving time with the GUI is significantly shorter than with the CUI interface. How can we explain this difference? Our first interpretation of this outcome was the supposed different amount of 'transparency' (Ulich, et al., 1991). One aspect of 'transparency' is 'feedback' (see Dix, et al., 1993, pp. 318-321).

Interesting is the fact, that the GUI supports the user with less "visual feedback" (FB = 66%, see Table 3) on average than the CUI (FB = 73%). This amount of FB of the CUI is caused by 22 small DCs with FB = 100%; the GUI has only 14 DCs with FB = 100%. The amount of functional feedback seems not to be related to the advantage of GUIs. There must be another reason.

The "interactive directness" is not quite different between both interfaces
(CUI: ID = 24.7% for AFIPs and 23.2% for DFIPs versus
GUI: ID = 22.5% for AFIPs and 25.5% for DFIPs, see Table 4).

Only the two measures of "flexibility" show an important difference between both interfaces
(CUI: DFA = 12.1 and DFD = 10.1 versus
GUI: DFA = 19.5 and DFD = 20.4, see Table 3).
We interpret this result to the effect that flexibility must exceed a threshold to be effective (DFD, DFA > 15).

4.2 *Results and discussion of experiment-II*

If our interpretation of the outcome of experiment-I is correct then we can not find a significant performance difference for dialog structures that remain under the assumed threshold of 15. To control the factor of feedback we carried out this second experiment with a multimedia information system that has 100% functional feedback for both interfaces (Brunner and Rauterberg, 1993).

We picked out a multimedia information system with a hierarchical dialog structure where DFA and DFD are clearly under 15. We implemented a comparable system with a net-shaped dialog structure where DFA and DFD have nearly the same ratio of flexibility as in experiment-I:
$DFA_{GUI} / DFA_{CUI} = 1.6$ and
$DFA_{MMnet} / DFA_{MMhier} = 1.2$;
$DFD_{GUI} / DFD_{CUI} = 2.0$ and
$DFD_{MMnet} / DFD_{MMhier} = 2.6$.

Table 2 Overview of our three empirical validation studies. DV means 'dependent variable' in the analysis of variance. The alpha-error is abbreviated with p.

Experiment	Interface type and dialog structure	Application type	Source of the empirical comparison study	Result of the empirical comparison study
I	CUI-hierarchical	Relational database	Rauterberg (1992)	DV: Task solving time
I	GUI-hierarchical	Relational database	Rauterberg (1992)	GUI << CUI ($p \leq .002$)
II	Multimedia-hierarchical	Information system	Brunner & Rauterberg (1993)	DV: Task solving time
II	Multimedia-net shaped	Information system	Brunner & Rauterberg (1993)	$MM_{hier} <= MM_{net}$ ($p \leq .085$)
III	CUI-hierarchical	Simulation tool	Grützmacher (1988)	DV: Target discrepancy
III	CUI-net shaped	Simulation tool	Grützmacher (1988)	$CUI_{hier} == CUI_{net}$ ($p \leq .784$)

Table 3 Comparison our three empirical validation studies relating to the quantitative measures ID, FB, DFA, and DFD. P is the number of all different dialog PATHs for an AFIP or a DFIP; D is the number of all different DCs.

Expe-riment	Interface type and dialog structure	P(AFIP)	ID(AFIP) %	P(DFIP)	ID(DFIP) %	D	FB %	DFA	DFD
I	CUI-hierarchical	434	24.7	362	23.2	36	73	12.1	10.1
I	GUI-hierarchical	547	22.5	570	25.5	28	66	19.5	20.4
II	Multimedia-hierarchical	241	25.1	34	28.1	68	100	3.6	0.5
II	Multimedia-net shaped	276	40.7	87	46.3	65	100	4.2	1.3
III	CUI-hierarchical	720	20.9	693	23.9	363	86	2.0	1.9
III	CUI-net shaped	490	15.8	1053	21.9	389	90	1.3	2.7

As we predicted, we can not find a significant performance difference between both types of dialog structures (see Table 2). To make sure that our results are not biased by our own expectations, we carried out a cross validation study. To do this, (1) we need the outcomes of an external independent comparison study between two different interfaces and (2) the possibility to apply our quantitative measures to all DCs of both interfaces. The empirical investigation of Grützmacher (1988) fulfils both conditions.

4.3 *Results and discussion of experiment-III*

The study of Grützmacher (1988) was carried out to investigate research questions in the context of how to control a complex domain with a simulation tool. One independent factor was varied: the dialog structure (hierarchical versus net-shaped). This simulation tool was implemented on a mainframe computer system with character oriented terminals (IBM 3270).

The dependent variable was not 'task solving time' but 'target discrepancy' as a performance measure. The sample consists of 20 users with the hierarchical dialog structure and 15 users with the net-shaped structure. The main result was that the factor 'dialog structure' was not significant. Given our interpretation of the last two experiments we expected a value for DFA and DFD under 15.

With the generous support of Grützmacher we were able to analyse all 752 dialog contexts for both interfaces. For the hierarchical CUI we get the following results: DFA = 2.0 and DFD = 1.9; for the net-shaped CUI: DFA = 1.3 and DFD = 2.7 (see last two rows in Table 3). These results for DFA and DFD of both CUI interfaces give us a sufficient evidence that the following assumptions are correct:

(1) We can measure the dialog flexibility in a task independent and quantitative way, and
(2) the values of DFA and DFD must exceed the threshold of 15.

5. CONCLUSION

Using the four quantitative measures for "feedback", "interactive directness" and "flexibility" to measure the interactive quality of user interfaces, we are able to classify the most common types: command, menu, desktop (see Rauterberg, 1993). The command interface is characterised by high interactive directness, but this interface type has a very low amount of visual feedback. Especially graphical interfaces (e.g., multimedia) can support users with sufficient interactive directness. GUIs are characterised by high dialog flexibility.

The presented approach to quantify usability attributes and the interactive quality of user interfaces is a first step in the right direction. The next step is a more detailed analysis of the relevant characteristics and validation of these characteristics in further empirical investigations. In the context of standardisation we can use our criteria to test user interfaces for conformity with standards.

ACKNOWLEDGEMENTS

We have to thank the following persons for their generous support: Dr. Karl Schlagenhauf and Raimund Mollenhauer at ADI Software Inc., Karlsruhe (D), Andreas Grützmacher and Prof. Eberhard Ulich at the ETH, Zürich (CH).

REFERENCES

Bevan, N., J. Kirakowski and J. Maissel (1991). What is Usability? In: *Human Aspects in Computing: Design and Use of Interactive Systems with Terminals* (H-J. Bullinger, (Ed.)), 651-655. Elsevier, Amsterdam.

Brunner, M. and M. Rauterberg (1993). *Hierarchische oder netzartige Dialogstruktur bei multimedialen Informationsystemen: eine experimentelle Ver-*

gleichsstudie. Technical Report MM-2-93. Institut für Arbeitspsychologie, Eidgenössische Technische Hochschule, Zürich.

Denert, E. (1977). Specification and design of dialogue systems with state diagrams. In: *International Computing Symposium 1977* (E. Morlet and D. Ribbens (Eds.)), 417-424. North-Holland, Amsterdam.

Dix, A., J. Finlay, G. Abowd and R. Beale (1993). *Human-Computer Interaction*. Prentice Hall, New York.

Grützmacher, B. (1988). *Datenpräsentation und Lösungsverhalten in einer komplexen, simulierten Problemsituation*. Unpublished Master Thesis. (Philosophische Fakultät I, Psychologisches Institut, Abteilung Angewandte Psychologie). Universität Zürich, Zürich.

IFIP (1981). *Report of the 1st Meeting of the European User Environment Subgroup of IFIP WF 6.5*. German National Center for Computer Science (GMD), P.O. 1316, D-5202 Sankt Augustin (Germany).

Jeffries, R. and H. Desurvire (1992). Usability testing vs. heuristic evaluation: was there a contest? *SIGCHI Bulletin* **24**(4), 39-41.

Karat, J. (1988). Software Evaluation Methodologies. In: *Handbook of Human-Computer Interaction* (M. Helander, (Ed.)), 891-903. Elsevier, Amsterdam.

Kirakowski, J. and M. Corbett (1990). Effective Methodology for the Study of HCI. In: *Human Factors in Information Technology Vol. 5* (H. Bullinger and P. Polson, (Eds.)). North-Holland, Amsterdam.

Laverson, A., K. Norman and B. Shneiderman (1987). An evaluation of jump-ahead technique in menu selection. *Behaviour and Information Technology* **6**(2), 97-108.

Paap, K. and R. Roske-Hofstrand (1988). Design of menus. In: *Handbook of Human-Computer Interaction* (M. Helander, (Ed.)), 205-235. Elsevier, Amsterdam.

Rauterberg, M. (1992). An empirical comparison of menu-selection (CUI) and desktop (GUI) computer programs carried out by beginners and experts. *Behaviour and Information Technology* **11**(4), 227-236.

Rauterberg, M. (1993). Quantitative Measures to Evaluate Human-Computer Interfaces. In: *Human-Computer Interaction: Applications and Case Studies* (M. Smith and G. Salvendy, (Eds.), Advances in Human Factors/Ergonomics Vol. 19A), 612-617. Elsevier, Amsterdam.

Rengger, R. (1991). Indicators of usability based on performance. In: *Human Aspects in Computing: Design and Use of Interactive Systems with Terminals* (H-J. Bullinger, (Ed.)), 656-660. Elsevier, Amsterdam.

Ulich, E., M. Rauterberg, T. Moll, T. Greutmann and O. Strohm (1991). Task orientation and user-oriented dialog design. *International Journal of Human-Computer Interaction* **3**(2), 117-144.

Wirth, N. and J. Gutknecht (1992). *Project Oberon - The design of an operating system and compiler*. Addison-Wesley, Reading (MA).

SUPPORTABILITY-BASED DESIGN RATIONALE

Guy A. Boy

European Institute of Cognitive Sciences and Engineering (EURISCO)
10, avenue Edouard Belin, BP 4032, 31055 Toulouse Cedex, France
Tel. +33-62 17 83 11, FAX +33-62 17 83 38, Email: boy@cert.fr

Abstract: How can we benefit from support people's experience in the design process, and how can we represent design rationale (DR) in order for it to be useful to a variety of end-users? This paper proposes a knowledge representation suitable for the rationalization of operational cognitive artifacts at design time. It enables the construction of cognitive functions that lead to the concept of software agents and active documents.

Keywords: Knowledge acquisition, knowledge representation, knowledge transfer, databases, design systems, documentation, human-centered design.

1. INTRODUCTION

Design rationale (DR) is represented by the reasons that lead to the design of engineered artifacts and the relationships between these artifacts. Most DR investigations focused on knowledge reuse for new designs (Potts & Bruns, 1988), i.e. from design to design. But almost nothing has been done on support issues as inputs to the DR construction. Users' requirements should be an integrated part of DR. Furthermore, requirements for the certification of modern aircraft are becoming more constraining for the manufacturers. Up to now, showing that a system works fine has been sufficient to certify it. New regulations will require the certification of the way this system was designed, i.e., instead of only certifying the performance of a product, constructors will have to explain how they have designed it also (SAE, 1994). This is another important reason to consider operations and support requirements at design time, and in particular in the construction of DR. There are two major questions: how can designers benefit from end-users' experience, and how can DR be represented in order for it to be useful to end-users? Support people, including operations, maintenance and training agents, spend enormous amounts of time trying to understand the functioning logic of the systems they are dealing with. We claim that appropriate viewpoints of DR improve understanding of functioning logic. The *supportability* of an engineered system is defined as the ability to support this system under a wider set of possible situations. Supportability depends on the traceability of DR seen from a support viewpoint. *Traceability* is the process of recording, maintaining and accessing DR. Our approach to this problem is to consider DR documentation as an intelligent assistant system (Boy, 1991a) which includes context-sensitive traceability.

Designers very rarely develop formal theories, but they do develop specific *artifacts* with very little explanation even if sometimes they try to make these artifacts as generic as possible. Artifacts are usually described and justified structurally and functionally by a text (including anecdotes or rational descriptions and explanations), graphics (e.g. a button for a screen interface), a video sequence, a simulation, or a straight piece of code. This paper introduces the description and justification of external mechanisms involved in the use of an artifact, in addition to the description and justification of internal mechanisms involved in the design of this artifact. One general answer to this problem is to construct *cognitive functions* as building blocks representing both designer knowledge and user knowledge involved in the artifact. The *block representation* (Boy & Caminel, 1989; Boy, 1989; Mathé, 1990; Boy, 1991ab) has been chosen to handle traceability problems. A block represents a *cognitive function* describing either how an artifact (internally) works and why, or how it is (externally) used and why.

Carrol and Rosson (1991) have already mentioned this trend by saying that a psychological DR is an enumeration of the psychological claims embodied by an artifact for the situations in which it is used. In large projects such as aircraft manufacturing, a new aircraft varies from a previous one because of the introduction of incremental modifications. Such modifications may result from either new technological possibilities or recommendations from the practitioners. Cognitive functions need to be identified from the description of operations episodes and adapted for new designs. In this framework, automation is seen as a transfer of *cognitive functions* from human agents to machine agents, e.g., in the past human pilots were controlling the basic flight parameters of the airplane manually (i.e., they developed tracking-task cognitive functions), since the 1930s an autopilot performs this cognitive function for them.

In this paper, supportability is stated as a main issue in the elicitation and storage of DR during the *life cycle* of an engineered system. In particular, end-users' experience feedback is an input to DR knowledge bases (Durstewitz, 1994). A distinction is made between stand-alone design and group design. For instance, when an aircraft is being designed, several viewpoints need to be kept in mind during the life cycle in order to reach a stable concept. A scenario describes the use of this approach to DR.

2. GROUP WORK DURING THE LIFE CYCLE OF COMPLEX SYSTEMS

This section advocates that during the life cycle of a complex system, there needs to be more cooperative activities and exchanges between the various agents. In particular, as there is usually little feedback between support people and designers, the former usually compensate design flaws and lack of documentation. The catch is that it is extremely difficult for the designers to anticipate support people's needs.

A distinction is made between *stand-alone design* (the activity of a regular designer) and *group design of a complex system* (design in a large project). The first kind of activity is *skill-based*, and the second one is strategic or *knowledge-based* in Rasmussen's sense (1986).

The first kind of design activity is very opportunistic. A study on design activity showed that even if a designer has a hierarchically structured plan for his activity, he uses it in an *opportunistic* way (Visser, 1989). Other design research mentions the opportunistic aspects of the design activity (Bisseret, Figeac-Letang & Falzon, 1988; Guidon, Krasner & Curtis, 1987; Kant, 1985; Ullman, Staufer & Dietterich, 1987; Whitefield, 1986).

Conversely large projects involve both a more classical planning activity and cooperative partnerships. Group design involves agents who interact with each others and exchange viewpoints. This interaction and exchange process is more or less effective according to the enterprise organization. These agents belong to various services such as designers, engineering, accounting, marketing, maintenance, training, operations, etc. The design cycle is usually sequential from design to operations, i.e., the overall design backbone is commonly organized into sequential phases. However, there are information exchanges between agents and subgroups of agents organized around viewpoints. They may interact through a shared database that is called enterprise memory or *technical memory*. From this perspective, design is a cooperative activity shared by various participants.

Lowry and Feaster (1987) have emphasized the life cycle cost (LCC) of a system. They have divided it into four phases: (1) the *mission definition* phase that involves the conceptualization of the system, i.e., definition of the problem to be solved and consideration of initial architectures; (2) the *design* phase including the design itself and the development and test of the prototype; (3) the *production* phase that entails the manufacturing of the product; (4) the *support* phase which involves training, actual use of the system, maintenance, repairs, etc. The life cycle costs (LCC) for a military or commercial system (Department Of Defense LCC Distribution) were provided by (Lowenstein & Winter, 1986). Results showed that the definition phase took less than 1% of the total LCC cost, the design phase less than 10%, the production phase about 30% and the support phase 60%. Furthermore, the Fiscal Year 1985 Congressional Budget Report gives a Space Shuttle LCC distribution where the support phase is 86% of the total LCC cost (only 6% for the design phase). These numbers clearly indicate the importance of support in the life cycle cost of a system. The main problem is that performance is currently almost the only major criterion taken into account during design and certification, with little or no emphasis on supportability. Lowry and Feaster have analyzed the 1986 Challenger accident. Among the essential issues they identified were accessibility, design criteria, integration, maintainability, management, procedure, reliability, design requirements, standards, training, and certification. All these issues show the current lack of *supportability*. Supportability includes maintenance, training and marketing issues. One important problem is to find the right balance between design for performance and design for supportability.

3. USER-ORIENTED DESIGN RATIONALE

3.1. The artifact-task-user triangle

An artifact is a physical or conceptual human-designed entity useful for a given class of users to perform specific tasks. Carroll and Rosson (1991) talk about transactions between tasks and artifacts in the *human-computer interaction* (HCI) world. It is sometimes very difficult to know if the task defines the artifact or if the artifact defines the task. In reality, users' profiles, tasks and artifacts are incrementally defined. The classical engineering tradition is centered on the construction of the artifact. The task and the user are usually taken into account implicitly. Tasks can be modeled from a *task analysis* or a model of the process that the artifact will help to perform. A specified task leads to a set of *information requirements* for the artifact. Conversely, the artifact sends back its own *technological limitations* according to the current availability of technology. Users can be incrementally taken into account in the design loop either through the development of syntaxo-semantic user models or through the adaptation of analogous user models. User modeling can be implicit or explicit, and leads to the definition of appropriate *user profiles*.

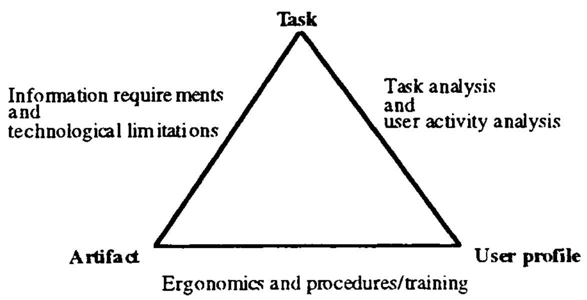


Figure 1. The artifact-task-user triangle

When a version of the artifact and the task are available, the user can use the artifact to perform the task. An *analysis of user activity* is then possible, and contributes to the modification of both the task and the artifact. The use of the artifact provides data to adapt both the artifact to the user (*ergonomics*), and the user to the artifact (*procedures* and *training*). The artifact-task-user triangle is described in Figure 1. It implicitly defines an incremental approach to design that is elsewhere described as a spiral model for software development (Boehm, 1988).

3.2. The four stage definition of design rationale

The following definitions are given in the framework of large projects. When an aircraft is being designed there are several decision levels that can be strategic, tactical or procedural. At the strategic level, there are several alternatives available, some of them are seen as possible by a restricted set of agents. In aeronautics, aviation requirements are defined at the strategic level. At this level, alternatives are called assumptions or beliefs. These assumptions evolve towards a stable solution through a process of merging several viewpoints, consensus and tests.

Justification of a smaller set of alternatives is performed by qualified personnel. These decisions may take time to be made. When alternative A* is developed, we are at the tactical and development levels, e.g., using a Computer-Aided Design (CAD) system. At these levels, design is more and more opportunistic. Designers tend to choose a solution straight away and iterate on it. They need to explain what they did afterwards. Note that before the CAD systems were available, design documentation was more incrementally and sequentially developed. This is because the design cycle was longer, and people had time to write memos and technical memoranda as outputs of design phases and inputs to the next ones. Today, engineers using CAD are experts in the development of stand-alone designs, but not necessarily in explaining their designs. This is why *explanation* has become a critical issue.

We will define *design rationale* as the reasons[1] R* why, at a given instant t, a project group decides to

[1] Design knowledge can be divided into five types (Freeman, 1990): *design data* that includes the structure, function, and properties of all artifacts used in the design; *unstructured design knowledge* that includes text or graphics; *structured design knowledge* that includes information in the form of tables, ratings, criteria, etc.; *refined design knowledge* that includes results of decision and classification methods like cluster or sensitivity analyses; and *expertise* that includes the engineering principles, models, design practices, and heuristics used in the design. All this information can be integrated in a multimedia database including text, graphics, and tables.

choose the best artifact alternative A* from a set of possible artifacts $\{A_1, A_2, ...\}$ according to some criteria. This definition departs from the ones given in the 1991 special issue of the Human-Computer Interaction Journal on DR (Carrol & Moran, 1991). Corresponding authors did not emphasize large projects nor group design.

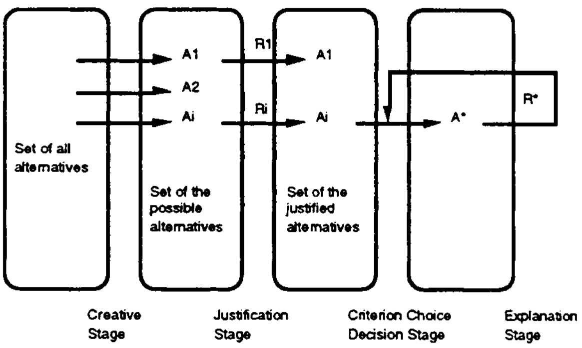


Fig. 2. Various stages in the construction of the design rationale.

The process of generating DR is divided into four stages (Figure 2). First, the project leaders must be aware of a set of possible alternatives $\{A_1, A_2, ...\}$. We call this stage the *creative stage*. Second, for each of these alternatives A_i, he has a "good" reason R_i to support that alternative, otherwise he will not include A_i in the set of possible alternatives. We call this stage the *justification stage*. Thus, alternatives and reasons for a given goal can be defined as a set of couples $\{(A_1, R_1), (A_2, R_2), ...\}$. A reason R_i is a function of the task for which the tool is being designed, the possibilities of the current technology, the time frame available to design and develop the tool, and other factors such as social and political constraints. During the justification stage, the expression of a reason R_j may indicate that the corresponding alternative A_j is too weak to be kept in the set of possible alternatives. Third, the designer has to chose among the alternatives and find a couple $(A^*=A_i, R_i)$ which is justified by the reason R*. To do this the designer has to chose a set of criteria that will rationalize his choice. We call this stage the *criteria choice/decision stage*. R* is the argument used by the designer to defend the decision that R_i is the best reason in the set $\{R_1, R_2, ...\}$. To define R*, it is necessary to define the method used to make the choice. Note any reason R_i of the set $\{R_1, R_2, ...\}$ may be modified during the decision process. Fourth, the expression of R* is called the *explanation stage*. Requirements follow from this explanation stage. The sequence [creative stage-justification stage-decision stage-explanation stage] is generally a closed-loop or iterative process. Furthermore, the more the domain is well mastered, the more the design loop may be open, i.e., very little or no interactions take place.

Currently, several DR investigators are focussing on the incremental modification of the selected alternative A* as well as on its associated explanation R* (i.e. the explanation stage). In large projects, we claim that each of these stages needs to be analyzed independently, and the mechanisms linking them within the design process need to be understood. At each stage, there are incremental modifications that may be caused by directives coming from users, designers or other people (sponsors in particular). Alternatives, justifications, criteria, decisions and/or explanations are stored at each stage of the DR construction cycle (Figure 2). Decisions may be made according to various viewpoints. In practice, viewpoints are characterized by actual services of the enterprise, e.g., engineering, financial support, marketing, maintenance, and so on.

4. A CASE STUDY

The goal of this exercise is to show how the supportability-based DR approach can improve the supportability of an engineered system. We provide a more detailed description of this case elsewhere (Boy & L'Ebraly, 1994).

The pilots of modern airplanes have become flight managers, i.e. they tend to supervise machine agents instead of maneuvering mechanical devices. For instance, the flight management and guidance system (FMGS) is a computer program (machine agent) that uses a navigation database including precompiled flight plans, and enables semi-automatic navigation. Before takeoff, the pilot inserts a flight plan into the FMGS. Consequently during the flight, the computer takes care of the guidance of the plane. The pilot can follow the evolution of the plane on a navigation display (ND) located in front of him. One of the current problems is the use of the multifunction command and display unit (MCDU) that is the user interface of the FMGS. Most pilots do not like to use the MCDU, e.g., they do not like its lateral location, its alphanumeric keyboard, its overloaded screen, its complex menu system, etc. EURISCO carried out a study to improve the current concept of MCDU which is briefly described below.

A cognitive analysis of the flight programming task enabled the elicitation of 8 primitive cognitive functions: browse, check, depress, enter, insert, modify, select and set. Each more complex cognitive function is a composition of these primitives. Each cognitive function has been represented in the form of a knowledge block. Blocks are appropriate for representing procedures. A block includes goals, a set of conditions and a set of actions to reach the goals. There are three types of conditions: contextual conditions (that are persistent preconditions), triggering preconditions and abnormal conditions (that can be preconditions or postconditions). Goals can be considered as normal postconditions. Selecting a specific action within a set of actions depends on the task, and its execution should lead to the satisfaction of a subgoal. Each specific action is attached to a set of abnormal conditions. Goals are suggested within a block either when its preconditions are satisfied or by another block.

Contextual conditions of a block are all the preconditions that vary slowly at use time. They can be hierarchically structured. In the case of a flight for instance, we can distinguish the following macroscopic contextual conditions: preflight, takeoff, climb, cruise, descent, and landing. The preflight context (or phase) can be divided into the following subcontexts: intialisation, flight plan input, and rolling. The transition from the *preflight* context to the *takeoff* context determines a new set of persistent parameters. Context is actually characterized by a set of persistent parameters or facts.

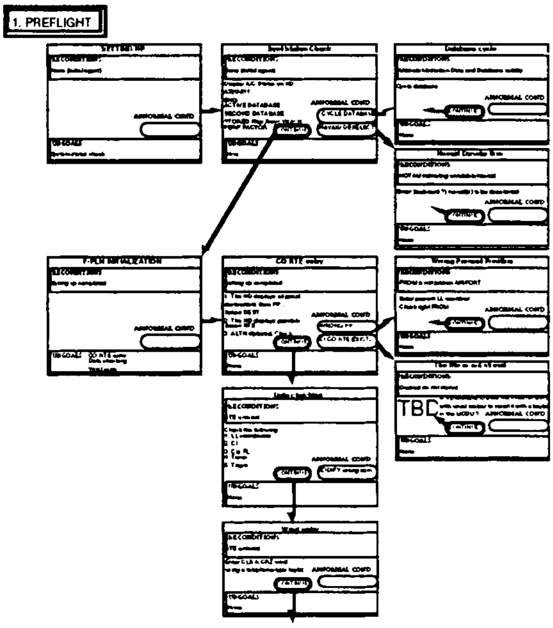


Fig. 3. Example of a network of blocks providing situated DR to both designers and support people.

A block knowledge base has been elicited from discussions with pilots and observations of simulated flights. An example of blocks is presented in Figure 3. Support people have been involved in this design experiment. In figure 3, each block represents a DR part that is meaningful for both designers and support people. In addition, each block is connected to an interface image of the real navigation display, i.e., access to DR is situated.

It was decided to remove the alphanumeric keyboard (*creative stage* with respect to the definitions provided in section 3), and to introduce a larger ND screen and a direct manipulation facility such as a trackball to manage information on this screen. This is an alternative to the current one. We have consequently developed a set of cognitive functions that are used by the pilots to program flight plans (*justification stage*). The efficiency of these functions was assessed using a GOMS analysis (part of the criterion for the *decision stage*) (Card *et al.*, 1983; Kieras, 1988). This analysis was performed with the help of operational people. It showed that the alternative to the current setup is much better, i.e., it takes much less time to access any FMGS function when one uses this alternative design (*explanation stage*). Furthermore, each cognitive function (knowledge block) was designed as a Hypercard card that gives the user the illusion of a "real" ND screen. In addition, each card is documented. Using such a hypermedia tool enables the generation of active documents that can both simulate real-world behavior and provide on-demand explanations.

Current results have shown that hypertext is a good programming tool for this kind of approach to design (Conklin & Burgess Yakemovic, 1991; Lee & Lai, 1991). We have already mentioned that hypertext systems increase accessibility, but they do not provide any built-in selectivity mechanism. For these reasons, the knowledge-based system technology can be very helpful for alleviating the selection problem and cognitive overhead of the user (both designers and operations people). Another reason for developing the block representation on top of hypertext systems is an easy mapping between the two: i.e., blocks are naturally associated with hypertext nodes containing DR.

This example has demonstrated that the supportability-based approach to rationalize design decisions can be extremely useful to explain why a part of a device has been specifically designed. This approach can be understood by designers, support people and certification people. Cognitive functions represented as knowledge blocks provide a powerful tool for mediating interaction between these practitioners. Blocks enable to incrementally store DR in context, and to easily retrieve and better interpret DR in context.

5. DISCUSSION AND FUTURE DIRECTIONS

Technical documentation was not produced in the same way in the past as it is produced now. Before the introduction of Computer Aided Design (CAD) tools, the design chain was very linear. Documents were produced sequentially to report (micro) decisions among alternatives. Documentation was developed sequentially and causally. Today, stand-alone designers sit facing a CAD screen, and iterate on an alternative that they find appropriate all day long. They do not try more alternatives than before, but they have the ability to correct flaws very easily. Today, practice shows that technical documentation is developed from explanations of CAD designs or simulations. CAD engineers seem to have difficulties in reporting what they did. This practice could be extremely beneficial when the end users are in the assessment loop, and are directly or indirectly involved in the documentation production. The notion of end-user is open and flexible; it includes operations, training, support, sale, etc. Designers can be end-users of reasons given by maintenance or training people, and conversely. This is a matter of viewpoint, e.g., from an aircraft designers' viewpoint pilots are end users.

The production of a technical documentation is a dynamic process. It is incremental by nature. A documentation is never finished. It is always annotated, modified and reconstructed. It has versions. Versioning must be handled by an appropriate mechanism. Again, versioning is a matter of viewpoint. Viewpoints can be generalized (merged) or specialized (spliced). Versioning for designers may not be relevant to training people who have their own versioning of the way they see an airplane for instance. An important factor is that technical documentation can be an output of design as well as an input. Assessment reports and requirements documents often have the same purpose: they assess the current design to improve it. Technical documentation is the formalization of current knowledge of the system being developed and used. This formalization can be performed by various agents. Each agent gives an opinion in his/her context. In this sense, documentation is context-sensitive according to the documentation developer. It is also context-sensitive according to the documentation user. If the user adds annotations, it becomes a documentation developer, and so on. We say that documentation is agent-based in the sense that each agent in contact with the documentation adds his/her view on it, and then modifies it accordingly. If viewpoints are kept in the documentation, context-sensitive mechanisms must be available to create and maintain these viewpoints.

The current practice of design engineering is often machine-centered: optimized for the tasks that can be performed by the machine itself, and neglecting to support the complementary role of the human in the loop (Boy & Gruber, 1990). The approach described in this paper takes an alternative view in which both design and operations are based on the utilization of DR documentation. Using such documentation increases the need to develop a context-sensitive indexing mechanism that is supported by the block representation. This mechanism is very similar to the mechanism used by end-users during operations. Operations procedures are developed from DR and may suggest modifications in DR (and by implication modification of the design itself). Procedures are based not only on DR, but also on users' requirements and suggestions when they are operating systems. We have extended the concept of "initial" DR to dynamic DR to capture the evolution of artifacts with time. In this view, the user should be a functional generator of DR. The more DR evolves, the more its traceability becomes a difficult issue. Evolution with time of DR and procedures to retrieve it (or to compile it, in the case of operations procedures) poses the problem of the formalization of context (contextual conditions). On one hand, the contextual conditions should be minimal to avoid excessive calculations. On the other hand, they must include as much information as possible to characterize the current situation. Definition of the contextual conditions is problem-dependent. It is very difficult, if not impossible, to define such contextual conditions *a priori*, since context comes from experience using the designed system and the corresponding documentation (compiled in the form of procedures or not). However, we have seen that contextual conditions can be incrementally augmented, i.e., expressions may be added to the current contextual conditions each time a new success (or failure) happens (Boy, 1991ab).

ACKNOWLEDGMENTS

This work was initiated when Dr. Boy was working at NASA Ames Research Center, California, USA. Thanks to Jody Gevins and Ana Cristina Bichara Garcia for many useful comments on very early attempts to this work. I am very grateful to Rolfe Folsom, David Lowry, Owen Green and Randy Fennel of Boeing for pertinent advice and knowledge on the practical aspects of design rationale and operations. Thanks to Hubert L'Ebraly who implemented the FMGS interface alternative prototype in cooperation with Airbus Industrie pilots and support people. Jean Pinet, Laurent Karsenty, Alfred Attipoe, Helen Wilson and Didier Lafrique also provided astute advice towards improving the quality of this paper.

REFERENCES

Bisseret, A., Figeac-Letang, C. & Falzon, P. (1988). *Modeling opportunistic reasoning: the cognitive activity of traffic signal setting technicians.* Research Report no. 898, INRIA, Roquencourt, France.

Boehm, B.W. (1988). A spiral model of software development and enhancement. *IEEE Computer*, Vol. 21, no. 5, May, pp. 61-72.

Boy, G.A. (1987). Operator Assistant Systems. *Int. J. Man-Machine Studies*, 27, pp. 541-554.

Boy, G.A. (1989). The Block Representation in Knowledge Acquisition for Computer Integrated Documentation. *Proceedings of the Fourth AAAI-Sponsored Knowledge Acquisition for Knowledge-Based Systems Workshop*, Banff, Canada, October 1-6.

Boy, G.A. (1991a). *Intelligent Assistant System*. Textbook. Published by Academic Press, London.

Boy, G.A. (1991b). *Computer Integrated Documentation*. NASA Technical Memorandum, Moffett Field, CA.

Boy, G.A. (1993). Knowledge acquisition in dynamic systems: how can logicism and situatedness go together? *Proceedings of the European Knowledge Acquisition Workshop*, Toulouse, France, Springer Verlag Publisher, Berlin.

Boy, G.A. & Caminel, T. (1989). Situation Patterns Acquisition improves the Control of Complex Dynamic Systems. *Proceedings of the Third European Knowledge Acquisition Workshop*, Paris.

Boy, G.A. & Gruber, T.R. (1990). Intelligent Assistant Systems: Support for Integrated Human-Machine Systems. *AAAI Spring Symposium on Knowledge-Based Human Computer Communication*, Stanford, March 27-29.

Boy, G.A. & L'Ebraly, H. (1994). *A task model and expertise representation for designing human-machine interaction devices - An aeronautical example*. (In French) EURISCO Report no. T-94-009-VI, October.

Card, S.K., Moran, T.P. & Newell, A. (1983). *The psychology of human-computer interaction*. Lawrence Erlbaum Ass., Pub., Hillsdale.

Carroll, J.M., Kellog,W.A. & Rosson, M.B. (1990). The Task Artifact. Research Report, IBM Research Division, T.J. Watson Research Center, Yorktown Heights, NY 10598. To appear in J.M. Carroll (Ed.), *Designing Interaction: Psychology at the Human-Computer Interface*. New York: Cambridge University Press, 1991.

Carroll, J.M. & Moran, T. (1991). Special Issue on Design Rationale. *Human-Computer Interaction*, Lawrence Erlbaum Associate, Inc.,Volume 6, pp. 281-318.

Carroll, J.M. & Rosson, M.B. (1991). Deliberated Evolution: Stalking the View Matcher in Design Space. *Human-Computer Interaction*, Lawrence Erlbaum Associate, Inc.,Volume 6, pp. 281-318.

Chen, H., & Dhar, V., (1987), "Reducing Indeterminism in Consultation: A Cognitive Model of User/Librarian Interactions", Proceedings of AAAI'87, Seattle, WA.

Conklin, J. (1987). Hypertext: An Introduction and Survey. *IEEE Computer*, September.

Conklin, J. & Burgess Yakemovic, K.C. (1991). A Process-Oriented Approach to Design Rationale. *Human-Computer Interaction*, Lawrence Erlbaum Associate, Inc.,Volume 6, pp. 357-392.

Durstewitz, M. (1994). Reuse of experience for industrial applications. *Proceedings of the First European Conference on Cognitive Science in Industry*, Luxembourg, September.

Fiscal Year 1985 Congressional Budget. (1985). *Cost Per Flight Operations Costs*. Office of Management and Budget.

Freeman, M.S. (1990). *Design knowledge capture in the Space Station Freedom Program*. SSFP Program Office, System Engineering and Analysis Group, NASA, Reston, VA. Presentation given February 21.

Guidon, R., Krasner, H. & Curtis, B. (1987). Breakdowns and processes during the early activities of software design by professionals. In G. Olson, S.S. Sheppard & E. Soloway (Eds.), *Empirical Studies of Programmers: Second Workshop*. Ablex Publishing Corporation, Norwood, NJ.

International Association for Nuclear Energy. (1986). Tchernobyl Nuclear Power Plant Accident and its Consequences. Working Document for the USSR National Committee on Nuclear Energy Use. Wien, Austria, August.

Kieras, D.E. (1988). Towards a practical GOMS model methodology user interface design. In M. Holander (Ed.), *Handbook of Human-Computer Interaction*. North-Holland Elsevier, Amsterdam, pp. 135-158.

Lee, J. & Lai, K.Y. (1991). What's in Design Rationale? *Human-Computer Interaction*, Lawrence Erlbaum Associate, Inc.,Volume 6, pp. 251-280.

Lowenstein, A. & Winter, G. (1986). VHDL's Impact on Test. *IEEE Design & Test*. NASA RECON 86A35709. April.

Lowry, D.J. & Feaster, T.A. (1987). Regarding Space Leadership through Control of Life Cycle Costs. *25th International Space Conference*, Cocoa Beach, Florida.

Mathé, N. (1990). *Assistance Intelligente au Controle de Processus: Application a la Telemanipulation Spatiale*. Ph.D. Dissertation, November, ENSAE, Toulouse, France.

Potts, C. & Bruns, G. (1988). Recording the Reasons for Design Decisions. IEEE Computer, pp. 418-427.

Rasmussen, J. (1986). *Information Processing and Human-Machine Interaction - An Approach to Cognitive Engineering*. North Holland Series in System Science and Engineering, A.P. Sage, Ed

Rogers, W.P. (1986). *Presidential Commission on the Space Shuttle Challenger Accident Report*. U.S. Government Printing Office: 688-274/58381, June.

SAE (1994). *Systems Integration Requirements.* ARP 4754, System Integration Requirements Task Group, AS-1C, ASD, Draft 43a, March 13, 1994.

System Requirements Document. Operations Analysis and Modeling Information Systems (OAMIS). Boeing Aerospace Operations Note. October 28, 1988.

Ullman, D., Staufer, L.A. & Dietterich, T.G. (1987). Towards expert CAD. *Computer in Mechanical Engineering*. 6, pp. 56-70.

Visser, W. (1989). More or Less Following a Plan during Design: Opportunistic deviations in Specification. *International Journal of Man-Machine Studies*, Special Issue on Empirical Studies of Programmers. Also INRIA Research Report, Roquencourt, BP 105, 78153 Le Chesnay Cedex, France.

Whitefield, A. (1986). An analysis and comparison of knowledge use in designing with and without CAD. In A. Smith (Ed.), *Knowledge Engineering and Computer Modeling in CAD. Proceedings of CAD'86*. Seventh International Conference on the Computer as a Design Tool. London: Butterworths.

A PROBABILISTIC METHODOLOGY FOR THE EVALUATION OF ALERTING SYSTEM PERFORMANCE

James K. Kuchar
R. John Hansman, Jr.

Department of Aeronautics and Astronautics
Massachusetts Institute of Technology
Cambridge, MA 02139

Abstract: A probabilistic methodology for evaluating hazard alerting systems is described that can be used in vehicle, transportation system, and process control applications. A means of showing the tradeoff between false alarms and missed detections is presented using signal detection theory concepts. The methodology accounts for uncertainties in measurement sources, alerting thresholds and displays, the human operator, and the situation dynamics. An example demonstration of the methodology is provided using the Traffic Alert and Collision Avoidance System (TCAS), an alerting system designed to prevent mid-air collisions between aircraft.

Keywords: Alarm systems, Human factors, Human supervisory control, Probabilistic models, Probabilistic risk assessment, Safety analysis, System analysis, System models

1. INTRODUCTION

A hazard alerting system is one of several safety components typically found in complex human-operated systems such as vehicles, traffic control systems, and process control applications. In these applications, the operator performs a task with the system using operating procedures and informational displays that provide feedback on the system and on the environment in which it operates (Sheridan, 1992; Wickens, 1992). Because unexpected hazards can be encountered, the alerting system serves to warn the operator that additional action may be needed to avoid an undesirable incident.

As the alerting system operates, it either remains silent or issues an alert that indicates that a hazard exists. Typically, this decision is based on whether certain states exceed critical values that form an alerting threshold. However, due to errors in measurements or limitations in design, faulty alerting decisions can occasionally occur. In particular, a system may fail to alert when necessary (termed a missed detection), or may issue an alert when one is not needed (termed a false alarm). Missed detections can result in an accident if the operator does not become aware of the hazard through other informational displays. False alarms have been shown to reduce the operator's confidence in the alerting system (DeCelles, 1991; Mellone and Frank, 1993). Alerting system performance is therefore dependent on the probabilities of missed detection and false alarm. Both types of error are undesirable and generally cannot be eliminated simultaneously. Rather, some tradeoff between false alarms and missed detections must be made.

A number of issues affect the ability of the alerting system to operate without missed detections and false alarms. Although many of these issues are common across applications, current design methods typically follow an *ad hoc*, evolutionary process. As alerting systems become more capable and more complex, however, the issues on which to focus design efforts become less evident. In particular, the effect of the human's response to an alert on the performance of the alerting system can be difficult to determine.

This paper presents a generalized model of alerting systems and a methodology for evaluating alerting system performance. By recasting the alerting decision as a signal detection problem, classical Signal Detection Theory methods can be used. Uncertainties in the human response and in

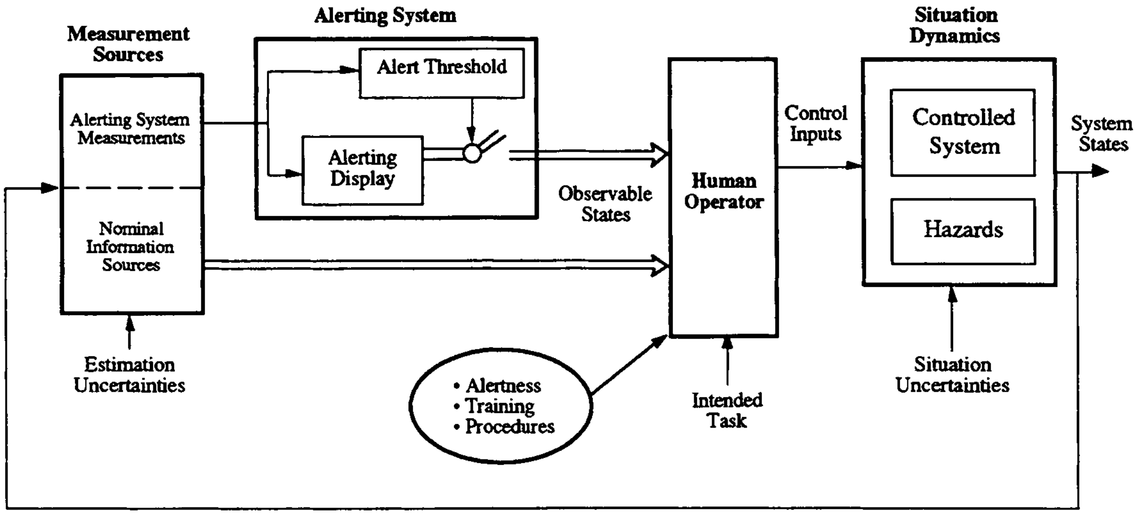


Fig. 1. Generalized Alerting System Model

measurements are considered. The methodology is then demonstrated through an example evaluation of the Traffic Alert and Collision Avoidance System (TCAS), used on civil jet transports.

2. ALERTING SYSTEM MODEL

Figure 1 shows a generalized control-system model of an alerting system that was developed by the authors. This model is based on a state-space representation of hazard situations. The state vector includes internal parameters such as a vehicle's spatial location, velocity, or acceleration, and external parameters describing a hazard's location, size, or severity. In this representation, there are four main elements in the control loop: Measurement Sources, Alerting System, Human Operator, and Situation Dynamics.

The Measurement Sources block represents the components that provide estimates of the system states to the human and to the alerting system. In general, these observable states are only a subset of the complete set of states that describe the situation.

The Alerting System uses a set of alerting thresholds to determine if an alert is warranted. When appropriate, alert information is provided to the Human Operator through aural or visual alerting displays.

The Human Operator uses information from the alerting system and other measurement sources to make necessary control inputs to the Situation Dynamics. The operator's response is also a function of the intended task and factors such as experience, fatigue, and training. These control inputs affect the evolving state vector, which is fed back to the Measurement Sources as the loop repeats.

Figure 2 shows an example situation in which the state vector is located in state-space near a hazard. Ideally, the alerting system should only alert the operator when it is clear that the hazard will be encountered unless action is taken. However, due to errors in the state measurements, the current state values are uncertain. In addition, errors in the extrapolation of the state trajectory (from an uncertain operator response, for example) result in a probabilistic future state trajectory. Thus, whether an alert is truly needed in a given situation is likewise uncertain, and missed detections or false alarms can occur. It is therefore necessary to consider the problem from a probabilistic standpoint to analyze the effects of uncertainties on the ability of the system to operate without missed detections and false alarms.

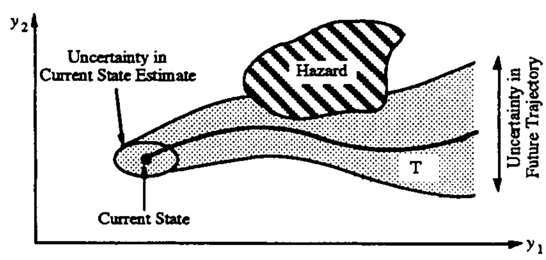


Fig. 2. Example Hazard Encounter Situation

3. PROBABILISTIC ANALYSIS METHOD

To determine if an alert is needed in a given situation, it is first necessary to find the probability that an encounter with a hazard will occur along a given trajectory, termed an *incident*, and denoted I. For example, in Figure 2 an alert may be necessary if the probability of an incident is high. If the probability of an incident is low, then an alert may not be needed.

Given a current state estimate, y, the probability that an incident will occur along a probabilistic trajectory T is written as $P_T(I \mid y)$. A generalized method for calculating $P_T(I \mid y)$ has been developed by Kuchar and Hansman (1995) and is outlined here.

The basic methodology for calculating $P_T(I \mid y)$ is shown in Figure 3. A set of equations that describe the dynamics of the situation and probability density functions (PDFs) describing the uncertainties in the parameters that define the trajectory are needed. For example, measurement uncertainties, the response time delay, and the aggressiveness of an avoidance maneuver can each be described by appropriate PDFs. The PDFs are obtained from hardware specifications or by a statistical analysis of actual or simulated hazard encounter situations.

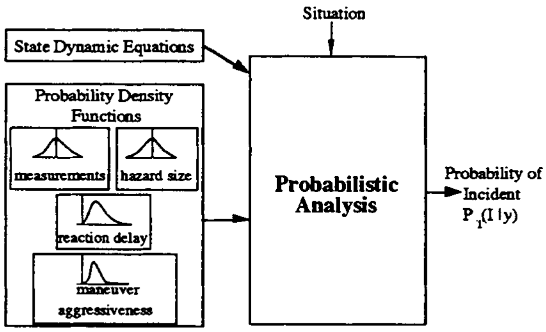


Fig. 3. Probabilistic Analysis Schematic

The procedure takes the equations of dynamics and the PDFs and, through numerical integration or Monte Carlo evaluation, produces the probability of an incident. The flexibility of the procedure lies in its ability to account for varied hazard types and to treat the PDFs as generic modules, allowing for a rapid comparison between the effects of different PDFs on $P_T(I \mid y)$.

3.1 Probabilistic Models

Because the measurements, the human response, and the system dynamics contain uncertainties, a probabilistic model of these components is needed to examine their effects on the overall alerting system. Measurement and dynamic uncertainties can generally be modeled based on hardware specifications. For example, a certain sensor may provide measurements that contain normally-distributed noise with a known variance.

The human's response to an alert is likewise uncertain. This uncertainty can generally be broken down into a probabilistic response latency and a set of random variables that specify the type of avoidance maneuver that is used. In an aeronautical application, for example, an avoidance maneuver to avoid terrain can be modeled as having a certain probabilistic aircraft pitch rate, bank angle, and thrust schedule.

3.2 Alerting Decision Outcomes

The alerting decision must balance the need to alert the operator sufficiently early that an incident can be avoided against the desire to only alert when absolutely necessary. To determine if an alert is warranted in a given situation, it is therefore necessary to examine the hypothetical outcomes of the alerting decision. This decision is analogous to the signal detection problem of determining if a signal is present in background noise (Swets and Pickett, 1982). In particular, if no alert is issued, the state continues along what is termed the projected *Nominal Trajectory*, denoted N. In response to an alert, there is, in general, a discrete change in the actions of the operator and the state follows a different trajectory that may avoid an incident, termed the *Avoidance Trajectory* and denoted A. Both N and A are, in general, probabilistic just as T is in Fig. 2.

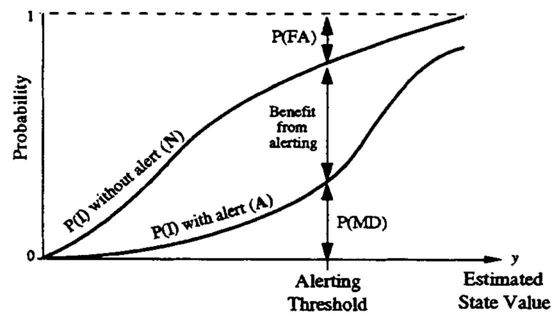


Fig. 4. Example Probability of an Incident Along Nominal and Avoidance Trajectories

An example plot of $P_N(I \mid y)$ and $P_A(I \mid y)$ is shown in Figure 4 as a function of an estimated state value, y. As shown, as the value of y increases, the probability of an incident occurring along N or A likewise increases. The probability that an incident would have been avoided without an alert is given by $1 - P_N(I \mid y)$, and is analogous to the probability of a false alarm, P(FA), in Signal Detection Theory. The probability that an incident will occur even though an alert is issued is given by $P_A(I \mid y)$, and is analogous to the probability of a missed detection, P(MD). One additional metric of system performance is often used: the probability of correct detection, P(CD), defined as P(CD) = 1 - P(MD). Also shown in Figure 4 is the benefit from alerting: the reduction in the probability of an incident that is possible because the alert is issued. As the threshold location is changed, there is a tradeoff between false alarms and missed detections. For example, moving the threshold to the left in Figure 4 increases the probability of false alarm, and therefore reduces the benefit from alerting.

3.3 System Operating Characteristic (SOC) Curves

The tradeoff between false alarms and missed detections can be viewed using a System Operating Characteristic (SOC) curve, similar to the Receiver Operating Characteristic (ROC) used in Signal Detection Theory (Swets & Pickett, 1982; Sheridan & Ferrell, 1974). SOC curves show the tradeoff between P(FA) and P(CD) as a function of the alerting threshold location, as shown in Figure 5. Each choice of an alerting threshold maps onto a single point along the SOC curve (two examples are shown in the figure).

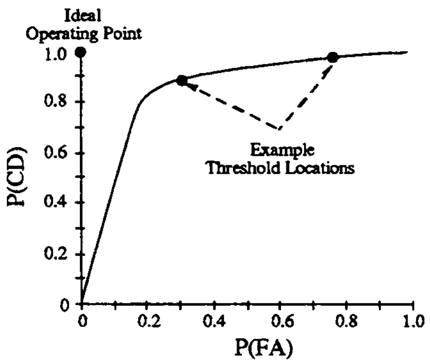


Fig. 5. Example System Operating Characteristic (SOC) Curve

Given a certain system design in terms of sensor accuracy, for example, the tradeoff between P(FA) and P(CD) is constrained to lie on a single SOC curve. P(FA) can then be balanced against P(CD) by changing the alerting threshold location, which changes the operating point on the SOC curve. Increasing sensor accuracy or improving the operator's response results in a shift in the SOC curve toward the upper-left ideal operating point.

Thus, given a definition of the system (which includes a probabilistic description of the human's response), the threshold location can be chosen according to the relative desirability of false alarms and correct detections. However, unless changes are made in the system design or in the operator's response, the system's performance will be constrained to follow a certain SOC curve.

4. EXAMPLE APPLICATION OF THE METHODOLOGY

The Traffic Alert and Collision Avoidance System (TCAS) is being used on transport aircraft in the U.S. to alert flight crews to potential mid-air collisions. This example addresses a situation in which an intruder is flying directly toward an aircraft equipped with TCAS (Figure 6). The intruder is currently above the TCAS aircraft and is projected to descend through the TCAS aircraft's altitude without leveling off. This situation has been known to produce false alarms in actual practice and provides an interesting example with which to apply the methodology.

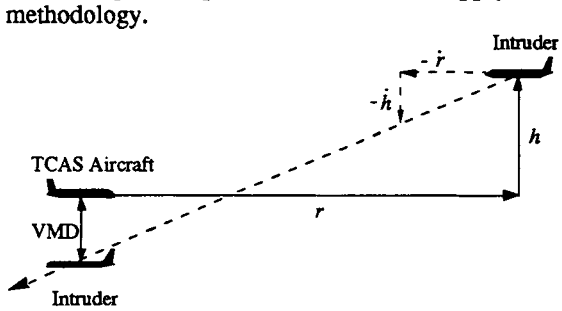


Fig. 6. Example Potential Collision Situation

Assuming straight-line extrapolations of each aircraft's trajectory, the Vertical Miss Distance (VMD) is a function of the range (r), range-rate ($\dot{r}$), relative altitude (h), and relative altitude-rate ($\dot{h}$):

$$\text{VMD} = h - \dot{h}\frac{r}{\dot{r}} \qquad (1)$$

TCAS uses a complex set of alerting thresholds based on estimates of the four parameters discussed above (range, range-rate, altitude, altitude-rate). A complete description of the TCAS alerting logic (called Version 6.04A) is provided in a set of operating specifications (RTCA, 1983; MITRE, 1993). In the situation used in this example, TCAS V6.04A issues an alert approximately 22 seconds before the projected time of impact.

When a TCAS alert is issued, the system determines whether a climb or a descent will provide the largest vertical separation between the aircraft. In the situation used here, if the intruder continues its descent, a climbing maneuver by the TCAS aircraft will provide the greatest separation between the aircraft. If, however, the intruder levels off after the alert is issued, a potential collision is induced by the alert as the TCAS aircraft climbs into the now-level intruder (Figure 7).

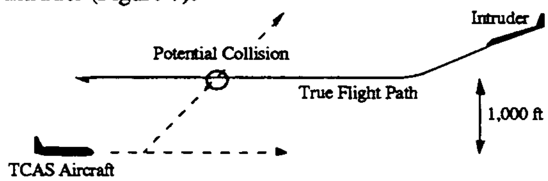


Fig. 7. Potential Induced Collision Due to Climb Alert

Therefore, TCAS must balance a desire to alert early and provide greater separation if the intruder continues to descend, and a desire to postpone alerting until the intruder's intentions are more well known.

4.1 Situation Parameters

The example described here assumes that the two aircraft are flying at an altitude of 4,500 m (15,000 ft) above Mean Sea Level, on a direct collision course at a closure rate of 200 m/s (400 knots). The intruder is descending toward the TCAS aircraft at 13 m/s (2,500 ft/min).

The intruder is assumed to continue its descent with probability 0.25 and to level off 305 m (1,000 ft) above the TCAS aircraft with probability 0.75. If the intruder levels off, it begins a maneuver such that a 0.1 g acceleration pull-up will result in a nominal separation between the aircraft of 305 m (1,000 ft). The actual altitude at which the level-off maneuver begins is modeled as a normal distribution with a standard deviation of 24 m (80 ft) (approximately 2 seconds of flight time).

If an alert is not issued, the TCAS aircraft is assumed to fly at a constant altitude (the Nominal Trajectory). In response to an alert, the TCAS aircraft flies a standard TCAS avoidance maneuver: 5 second delay followed by a 0.25 g acceleration pull-up maneuver to a climb rate of 7.6 m/s (1,500 ft/min) (RTCA, 1983). This climbing maneuver defines the Avoidance Trajectory.

A Monte Carlo simulation of the TCAS parameter estimation logic was performed to obtain the steady-state standard deviations of each state estimate. Approximate standard deviations on the parameters for this situation are shown in Table 1. Normally distributed probability density functions were created with the standard deviations shown in Table 1 to model the uncertainty in state estimates.

Table 1 TCAS Parameter Estimate Accuracies

Parameter	Estimate Standard Deviation
r	5.5 m (18 ft)
$\dot{r}$	1.8 m/s (6 ft/s)
h	25 m (83 ft)
$\dot{h}$	0.17 m/s (0.6 ft/s)

4.2 Human Response Model

The human response is assumed to include a probabilistic response delay, τ, followed by a 0.25 g pull-up maneuver to a 7.6 m/s (1,500 ft/min) climb rate. The response delay is modeled as a Gamma Distribution (Figure 8), which is a smooth, skewed distribution similar to that which might be expected (Hogg & Tanis, 1988). This PDF also provides a non-zero probability of very long response times representative of a flight crew that disregards or misunderstands an alert. A more detailed study of TCAS could use statistics from actual events or flight simulations to build other PDFs of response time delay.

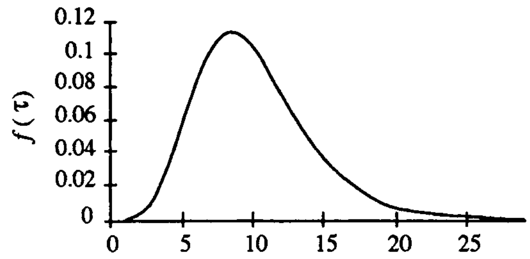


Fig. 8. Example Gamma Distribution $\alpha = 7$, $\theta = 1.4$

The Gamma Distribution is given by:

$$f_{\tau}(\tau) = \frac{1}{\Gamma(\alpha)\,\theta^{\alpha}}\,\tau^{\alpha-1}\,e^{-\tau/\theta} \qquad (2)$$

where $\Gamma(\alpha) = (\alpha - 1)!$

The mean, $\bar{\tau}$, of the Gamma Distribution is given by:

$$\bar{\tau} = \alpha\theta \qquad (3)$$

The PDFs used here have a parameter value of $\alpha = 7$ and θ is varied to provide a desired mean value.

4.3 Calculation of the Probability of an Incident

The PDFs used in this example, representing uncertainties in range, range-rate, altitude, altitude-rate, and response latency, were used to calculate the probability of an incident along the Nominal and Avoidance Trajectories. These PDFs were numerically integrated to determine the probability that an incident would occur along a given trajectory. This integration can be summarized by Equation (4).

$$\mathrm{P(I)} = \int_{All\ PDFs} G(r,\dot{r},h,\dot{h},\tau)f(r)f(\dot{r})f(h)f(\dot{h})f(\tau) \qquad (4)$$

where, for example, $f(r)$ is the PDF for r, and G is a function that takes a value of 1 for combinations of parameters that result in a collision, and 0 otherwise. A collision is defined to occur if the Vertical Miss Distance is less than 30.5 m (100 ft).

4.4 Results

An SOC curve for this example is shown in Figure 9 for the case in which the mean response delay is 5 seconds. Note that the SOC curve for this example appears quite different from conventional ROC curves used in Signal Detection Theory, but the SOC curve can still be used in the same manner. As shown, the current alerting threshold location (TCAS V6.04A) is such that a near-minimum of false alarms are expected. Any further changes to reduce P(FA) will result in a rapid increase in P(MD). Due to the large uncertainties in the altitude of the intruder, the false alarm probability remains above approximately 0.4 regardless of the threshold location.

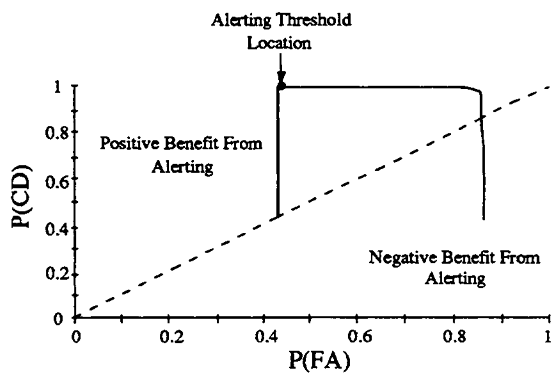


Fig. 9. System Operating Characteristic (SOC) Curve for TCAS Example ($\bar{\tau}$ = 5 sec).

Only by increasing the breadth or accuracy of information available to TCAS could the false alarm probability be reduced. There is also a region on the SOC curve where P(CD) is less than P(FA): it is more dangerous to alert than to not alert. This region corresponds to alerting thresholds that are set such that there is a greater probability of climbing into a leveled intruder than colliding with an intruder that continues its descent.

Figure 10 shows the probability of an incident when an alert is issued, P(MD), at the V6.04A threshold location as a function of the mean response latency, $\bar{\tau}$. A mean response delay of 5 sec results in a value of P(MD) of approximately 2×10^{-5}. An increase in the mean response time to 7 seconds results in an increase in the value of P(MD) to approximately 5×10^{-4}. Thus, an additional two second delay over that assumed by TCAS logic reduces the safety of the system by a factor of 25.

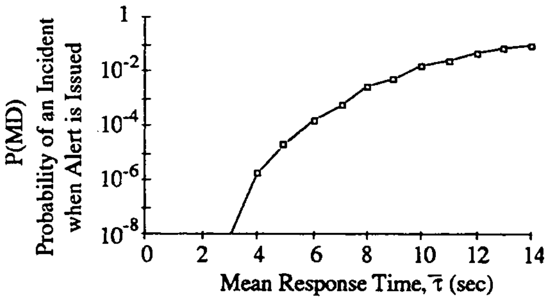


Fig. 10. Effect of Mean Response Time on Probability of an Incident.

5. CONCLUSION

A methodology for modeling and evaluating hazard alerting systems has been developed. This methodology accounts for uncertainties in measurements, the human response to an alert, and situation dynamics. The alerting decision is shown to be analogous to problems in Signal Detection Theory. This connection enables the use of conventional analysis tools to examine the tradeoffs between false alarms and missed detections. Parametric studies are also possible to examine the effect of the human response on alerting system performance.

The methodology is demonstrated in an application using the Traffic Alert and Collision Avoidance System (TCAS) in use on U.S. jet transport aircraft. For the situation covered here, the current alerting threshold is shown to be located to effectively minimize false alarms before a rapid increase in missed detections occurs. The sensitivity of the alerting system's performance is also shown as a function of the mean response latency. This example serves to illustrate the basic concepts of using probabilistic analysis to evaluate alerting systems. A more detailed examination is certainly possible, and will be necessary in an actual design situation.

The methodology is presented here using an aeronautical example. However, the methodology is easily extended to any alerting system application, including other vehicle types, large-scale transportation systems, and process control or medical applications.

ACKNOWLEDGMENT

This work was supported by the Federal Aviation Administration under Grant #92-6-0001, and by the Boeing Commercial Airplane Company.

REFERENCES

DeCelles, J. L. (1991). The delayed GPWS response syndrome. Aviation Research & Education Foundation. Herndon, VA. July.

Hogg, R. V., and E. A. Tanis (1988). *Probability and Statistical Inference*. Macmillan. New York.

Mellone, V. J. and S. M. Frank (1993). Behavioral impact of TCAS II on the national air traffic control system. *Seventh International Symposium on Aviation Psychology*. The Ohio State University. April 27.

The MITRE Corporation (MITRE) (1993). *Final V6.04A Change Pages*. Courtesy of Andrew D. Zeitlin. September 17.

Kuchar, J. K. and R. J. Hansman (1995). A Unified Methodology for the Evaluation of Hazard Alerting Systems. MIT Aeronautical Systems Laboratory Report No. ASL-95-1.

Radio Technical Committee on Aeronautics (RTCA) (1983). *Minimum Performance Specifications for TCAS Airborne Equipment*. Document No. RTCA/DO-185. Washington, D.C.

Sheridan, T. B. and W. R. Ferrell (1974). *Man-Machine Systems: Information, Control, and Decision Models of Human Performance*. MIT Press. Cambridge, MA.

Sheridan, T. B. (1992). *Telerobotics, Automation, and Human Supervisory Control*. MIT Press. Cambridge, MA.

Swets, J. A. and R. M. Pickett (1982). *Evaluation of Diagnostic Systems*. Academic Press. New York.

Wickens, C. D. (1992). *Engineering Psychology and Human Performance*. Harper Collins.

APPLICATION OF THE ANALYTIC HIERARCHY PROCESS FOR MAKING SUBJECTIVE COMPARISONS BETWEEN MULTIPLE AUTOMATION/DISPLAY OPTIONS

Lee C. Yang
R. John Hansman, Jr.

Department of Aeronautics and Astronautics
Massachusetts Institute of Technology
Cambridge, MA 02139

Abstract: This paper describes the application of the Analytic Hierarchy Process (AHP) as an experimental tool for obtaining subjective preferences in Human Factors studies. The AHP utilizes redundant paired comparisons to obtain a relative ranking between alternatives under study. Compared with other approaches of analyzing subjective data, the AHP has the advantage of allowing multiple alternatives while accommodating a small sample size. It incorporates the use of judgment matrices, ratio scales, and eigenvector weightings in computing a weighted ranking of the options under study. The methodology is explained and demonstrated using data from a study involving cockpit displays of Enhanced Vision Systems.

Keywords: Human Factors, Decision Making, Comparing Elements, Man/Machine Interfaces, Information Analysis

1. INTRODUCTION

In the development and evaluation of Man-Machine interfaces, it is often desirable to obtain subjective preferences of multiple design options. When comparing various designs, the experimenter is often faced with the task of evaluating a multitude of subjective data to determine the preferred choice. In typical studies comparing automation or display designs, more than two options are often under evaluation making it difficult to produce statistically significant results from traditional approaches involving non-parametric rankings. These methods are further hindered in utility by the small subject groups often encountered in preliminary prototyping studies.

Originally developed by Saaty (1980) as a decision making aid, the Analytic Hierarchy Process (AHP) has found wide usage in a broad range of areas including market forecasting, architectural design, strategic policy planning, conflict analysis, project evaluations, and medical decision making (Vargas, 1990). It has also found interest as a possible tool for subjective workload assessment (Vidulich, 1989). In this paper, the AHP is described as an experimental tool for obtaining subjective preferences in Human Factors studies. It provides a means of evaluating multiple design options using weighted ranking scales. The methodology is demonstrated for a recently completed cockpit display study involving Enhanced Vision Systems.

2. METHODOLOGY

The Analytic Hierarchy Process is a decision aiding method developed by Saaty (1980) based on the use of relative judgment matrices. It breaks up multiple options into a series of paired comparisons which are then recombined to produce an overall weighted ranking. Typical techniques based on rankings involve assigning integer value ranks (of 1, 2, 3, etc.) to each of the different options under evaluation. Inherent in this approach is the loss of information regarding the relative size of the intervals between the rankings. However, this information can be a very important subjective measure and is a major premise upon which the AHP is based. In evaluating three different cockpit displays, for example, if one of the displays were radically different from the other two, assigning integer rankings to the displays may not be sufficient to express the subjective opinions of the subjects evaluating the displays. Some knowledge of the relative differences between the rankings is needed. Display A may be minutely better than Display B, but Display C may be overwhelmingly better than display A. It is precisely this type of information which the AHP methodology attempts to capture and utilize in its analysis of formulating a weighted ranking scale.

2.1 Data Collection

In the AHP, data is collected using a series of head-to-head paired comparisons between each pair of design options being considered. A typical scale in obtaining such data is shown in Figure 1 where Display A is being compared to Display B. Saaty (1980) and Mitta (1993) suggest a format showing the two items, or options, to be compared on opposite ends of a 17 slot rating scale. The scale is a measure of *dominance* of one alternative over the other. It uses five descriptors in a pre-defined order and allows a single space between each one for compromises. The descriptors suggested by Saaty (1980) and Mitta (1993) are "equal", "weak", "strong", "very strong", and "absolute". These wordings maybe modified to provide easier comprehension for a particular subject group. The scale allows the subjects to indicate their judgment regarding the *degree of dominance* of one option over the other. The subjects indicate not only that one alternative dominates the other, but the degree to which it dominates. The dominance scale may be referenced to metrics such as workload, ease of use, aesthetics, or overall preference depending on objectives of the study. In the example shown in Figure 1, the measure is simply the "better" display.

Each m_{ij} entry of **M** reflects the dominance of A_i (alternative i) over A_j (alternative j) as specified by the following scale:

$m_{ij} = 1$ if alternative i and j are of equal strength
$m_{ij} = 3$ if alternative i weakly dominates alternative j
$m_{ij} = 5$ if alternative i strongly dominates alternative j
$m_{ij} = 7$ if alternative i very strongly dominates alternative j
$m_{ij} = 9$ if alternative i absolutely dominates alternative j

Scale values 2, 4, 6, and 8 are used to reflect the compromise ratings in-between.

Notice that **M** follows a reciprocal structure with the following specifications:

i) $m_{ji} = 1 / m_{ij}$, for $m_{ij} \neq 0$
ii) $m_{ij} = 1$, for $i = j$ and $i, j = 1, 2, \ldots, n$

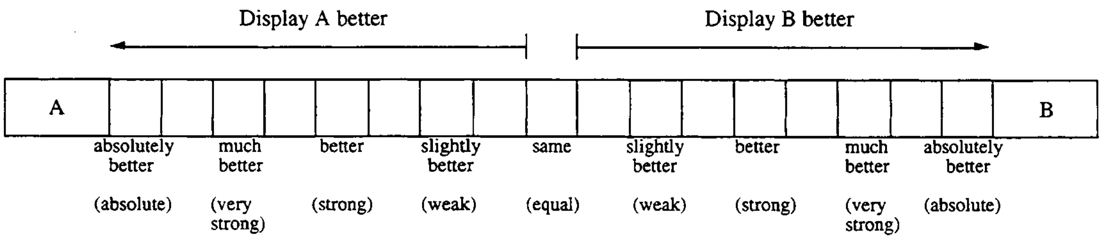


Fig. 1. Dominance Scale Used for Paired Comparison of Two Display Options

Given n items to compare, all possible pairs of alternatives must be considered. Thus each subject must make $n(n-1)/2$ comparisons. If three options were to be compared (n=3), then each subject would make 3(3-1)/2 = 3 paired comparisons. If there were four options (n=4), then six paired comparison would be made by each subject. In breaking up the analysis into a series of head-to-head comparisons, the results will indicate how each option compares relative to one another. Notice that by making all possible $n(n-1)/2$ paired comparisons, redundant measurements are available.

2.2 *Judgment Matrices, Ratio Scales, and Eigenvector Weightings*

For each subject, the data from the paired comparisons are placed in a judgment matrix **M** where the rows and columns represent the compared items. Each cell in the matrix represents the comparison of the item in that row to the item in that column. If n is the number of items, or alternatives, to be compared, then **M** is matrix of size $n \times n$ of the following form:

$$\mathbf{M} = \begin{array}{c} \\ A_1 \\ A_2 \\ \vdots \\ A_n \end{array} \begin{array}{c} \begin{array}{cccc} A_1 & A_2 & \cdots & A_n \end{array} \\ \begin{pmatrix} 1 & m_{12} & \cdots & m_{1n} \\ 1/m_{12} & 1 & \cdots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 1/m_{1n} & 1/m_{2n} & \cdots & 1 \end{pmatrix} \end{array} \qquad (1)$$

The reciprocal format of **M** comes from the AHP axiom of *ratio scales:* if A is x times more dominant than B, then B is $1/x$ the dominance of A. Furthermore, if C is y times more dominant than A, then C would be expected to be xy times more dominant than B. Note that the slots along the main diagonal, from upper left to lower right of the matrix, represent each item compared to itself, and therefore contain the value 1 to imply equality.

The next step is to produce an overall rating scale from all the pair-wise comparisons. Several methods have been proposed in the past to produce this rating scale, including eigenvector weighting, logarithmic least squares, and simple averages (Saaty, 1980, 1988, 1990a, 1990b; Vidulich, 1989; Budescu et al., 1986). This paper will concentrate on the eigenvector method used in the original AHP.

In the eigenvector method, the scale is produced by calculating the principal eigenvector of **M**. The principal eigenvector, $\mathbf{w} = [w_1\ w_2\ \cdots\ w_n]^T$, corresponds to the largest eigenvalue, λ_{max}, and can be determined from solving the following system of equations:

$$\mathbf{Mw} = \lambda \mathbf{w} \qquad (2)$$

According to the AHP methodology, **w** is typically normalized such that its components sum up to 1.

$$\sum_{i=1}^{n} w_i = 1 \qquad (3)$$

Note that this differs from the more common practice of normalizing the components to a magnitude of 1. Afterwards, the eigenvector can be referred to as the *priority vector*. **w** can be thought of as a vector in the space of the n different options where the magnitude of the components in each direction is a measure of the strength of the respective option. The degree of dominance between two options is then the ratio of their **w** components.

A maximum of n distinct eigenvalue-eigenvector pairs can be obtained from a $n \times n$ **M** matrix, but it is the eigenvector corresponding to the largest positive eigenvalue that is important. The maximum eigenvalue, λ_{max}, can be thought of as a measure of inconsistency in the paired ratings from the expected AHP ratio scale axiom. It can be shown that $\lambda_{max} \geq n$ (Saaty, 1990a). A $n \times n$ matrix, **M**, is said to be perfectly consistent if and only if $\lambda_{max} = n$. When this occurs, each column of **M** is a constant multiple of another column and there will be only one non-zero eigenvalue. As **M** deviates from this perfect consistency condition, λ_{max} will increase from the value n. Therefore, the ratio λ_{max}/n is an indication of consistency with $\lambda_{max}/n = 1$, indicating perfect consistency and higher values indicating the degree of inconsistency.

Example. A previous experiment conducted by Yang and Hansman (1994) will be used as an example. The objective was to evaluate three types of Enhanced Vision System displays in terms of both performance and subjective metrics. The AHP methodology was used to rank the three displays (A, B, C) in terms of overall subjective preference. The experiment was conducted on a part-task simulator with airline pilots using the prototype display formats to perform low visibility Category III landings. After completion of all the experimental runs and having flown with all three displays, the pilot were asked to make pair-wise comparisons using questionnaires similar to Figure 1. The experimental scenario required each individual to complete three paired comparisons. For each comparison, the subject indicated which one of the two displays was better and specify the factor by which it was better than its counterpart. The ultimate objective was to rank the three displays in terms of the overall "better" display.

The results of one pilot's judgments are tabulated in the following example matrix, **M**, below:

$$\mathbf{M} = \begin{pmatrix} 1 & 1/6 & 1/8 \\ 6 & 1 & 1/3 \\ 8 & 3 & 1 \end{pmatrix} \tag{4}$$

The entries m_{ij} represent the dominance of display i over display j in terms of the subjective preference. Determination of the eigenvalues and eigenvectors provided the following results:

$$\begin{aligned} \lambda_{max} &= 3.0735 \\ \mathbf{w} &= \begin{bmatrix} 0.0623 & 0.2851 & 0.6526 \end{bmatrix}^T \\ &= \begin{bmatrix} 0.0623 \\ 0.2851 \\ 0.6526 \end{bmatrix} \begin{matrix} \text{display A} \\ \text{display B} \\ \text{display C} \end{matrix} \\ \lambda_{max}/n &= 1.0245 \end{aligned} \tag{5}$$

The results indicated that display C was the most preferred by this particular subject, exceeding display A and display B by factors of 10.5 and 2.3, respectively. Display B ranked second and exceeded display A by a factor of 4.6. The value of $\lambda_{max}/n = 1.0245$ is close to unity, indicating good consistency.

2.3 Multiple Subjects

If data is obtained from more than one subject, a separate **M** matrix is tabulated for each person and separate priority vectors are computed. The matrix **W** is constructed from the eigenvectors $\mathbf{w}^k$, where k denotes the k-th subject. If the number of subjects is s and there are n options being compared, **W** would be a $n \times s$ matrix:

$$\mathbf{W} = (\mathbf{w}^1, \mathbf{w}^2, \cdots, \mathbf{w}^s) \tag{6}$$

W is called the individual ranking matrix because it contains the rankings of each individual. The columns of **W** are the normalized principle eigenvectors from each subject. Care must be taken to insure that all the **w**'s are normalized to sum up to 1.0. In other words, the columns of **W** must each add up to unity.

Example. Continuing with the experiment by Yang and Hansman (1994), six commercial airline pilots were used as test subjects. The **W** matrix from the study is shown below:

$$\begin{aligned} \mathbf{W} &= (\mathbf{w}^1, \mathbf{w}^2, \mathbf{w}^3, \mathbf{w}^4, \mathbf{w}^5, \mathbf{w}^6) \\ &= \begin{pmatrix} 0.0456 & 0.0526 & 0.1734 & 0.0510 & 0.0623 & 0.0457 \\ 0.1986 & 0.4737 & 0.7720 & 0.2270 & 0.2851 & 0.7870 \\ 0.7558 & 0.4737 & 0.0546 & 0.7220 & 0.6526 & 0.1673 \end{pmatrix} \end{aligned} \tag{7}$$

The columns of **W** were formed from the priority vectors of each **M** matrix, one for each of the six subjects. Notice that the components of each vector, $\mathbf{w}^k$, add up to 1.

2.4 Ranking of Subjects

The next step is to combine the priority vector of all the subjects into one final ratio scale. The AHP allows for results of each subject to contribute equivalently or differently depending on their skill level, experience, judgment ability, and so forth. Mitta (1993) suggests that the experimenter observe all the subjects during the trials and compare their abilities to make sound judgments. The subjects are ranked by the experimenter in a similar manner that the alternatives were ranked by the subjects to obtain the **M** matrix. With s subjects, the experimenter makes $s(s-1)/2$ paired comparison of the subjects and formulates a $s \times s$ matrix $\mathbf{M}'$, where the entries of $\mathbf{M}'$ reflect the factors by which one subject dominates another in judgment skill. The ranking of the subjects is available from the normalized principal eigenvector of $\mathbf{M}'$ and can be defined as $\mathbf{s} = [s_1 \; s_2 \; \cdots \; s_s]^T$. The system of equations may be written as:

$$\mathbf{M}'\mathbf{s} = \lambda_{max}\mathbf{s} \tag{8}$$

In most cases, however, the subjects should be given equal consideration in the final analysis; otherwise, the outcome could easily be biased toward a certain result. Given this, the contribution from all subjects are considered equally important and $\mathbf{M}'$ becomes a $s \times s$ matrix filled with all ones. Consequently, all entries of its principal eigenvector **s** will be the same; making the subjects all equally weighted, per se. It can then be shown that after normalizing, $\mathbf{s} = [1/s \; 1/s \; \cdots \; 1/s]^T$ where s is the number of subjects participating.

Example. Contribution from all six pilots ($s = 6$) were given equal consideration, thus $\mathbf{M}'$ was a 6×6 square matrix filled with all ones. The normalized principle eigenvector of $\mathbf{M}'$ was obtained from $\mathbf{s} = [1/s \; 1/s \; \cdots \; 1/s]^T$.

$$\mathbf{M}' = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix},$$

$$\mathbf{s} = \begin{bmatrix} 1/6 \\ 1/6 \\ 1/6 \\ 1/6 \\ 1/6 \\ 1/6 \end{bmatrix} \begin{matrix} \text{subject 1} \\ \text{subject 2} \\ \text{subject 3} \\ \text{subject 4} \\ \text{subject 5} \\ \text{subject 6} \end{matrix} \qquad (9)$$

2.5 Final Overall Ranking

The individual priority vectors from matrix **W** are algebraically combined to the ranking of the subjects, **s**, to obtain the final overall ranking. This final ranking is specified by a vector **r**, where **r** is the computed from the matrix multiplication:

$$\mathbf{r} = \mathbf{W} \cdot \mathbf{s}$$

r should be normalized to sum up to 1.0 if it is not already in that form. The components of $\mathbf{r} = [r_1 \; r_2 \; \cdots \; r_n]^T$ provide a weighted ranking scale of the n alternatives; the larger the value, the higher the rank. A rating of $r_1 = r_2 = \cdots = r_n$ would imply that all the items or alternatives were considered equal overall by the subjects. The advantage of this type of scale is the information provided about the relative differences between the rankings. A and B may both be preferred over C, but knowing that A is overwhelmingly preferred while B is only slightly preferred can be an important asset to the experimenter or design engineer. When looking at the relative differences, recall the AHP axiom of ratio scales used in developing the judgment matrix. The degree of dominance between two alternatives can be obtained by taking the ratio of their two values, r_i and r_j. Once this is done, it is useful to reference this ratio back to the qualitative descriptions used to form the scale of Figure 1. Taking r_i as the higher rated alternative, Table 1 can be used to make the conversion.

Table 1. Conversion of Ratio Scale to Qualitative Description

r_i / r_j	Dominance of r_i over r_j
1	equal
3	weak dominance
5	strong dominance
7	very strong dominance
≥9	absolute dominance

Example. Taking **W** and **s** from the ongoing example, the final overall ranking **r** can be easily computed to be:

$$\mathbf{r} = \mathbf{W} \cdot \mathbf{s} = \begin{bmatrix} 0.0718 \\ 0.4572 \\ 0.4710 \end{bmatrix} \begin{matrix} \text{display A} \\ \text{display B} \\ \text{display C} \end{matrix}$$

The results indicated that displays B and C were the most preferred by this particular subject group, exceeding display A by factors of 6.4 and 6.6, respectively. With the help of Table 1, this would imply "strong" to "very strong" preference over display A. Although display C had a slightly larger value than display B, they can be considered equal since the dominance ratio between the two was approximately 1.0.

The final data can be presented in a pie chart since the components of **r** sum to 1. An example is shown is Figure 2. The larger the area, the more dominant the display. The relative dominance between two displays can be obtained by taking the ratio of their respective areas.

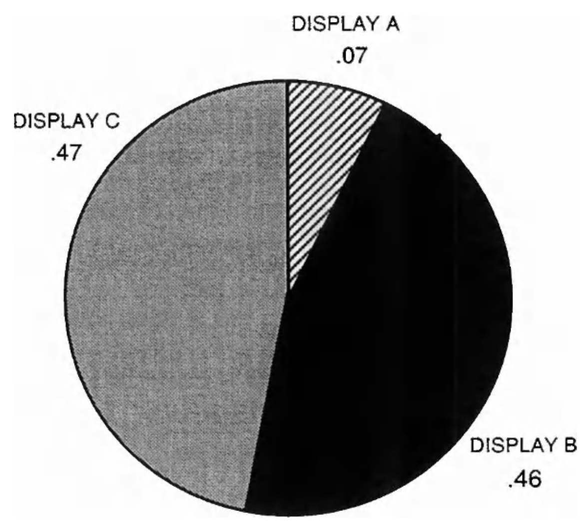


Figure 2. Final Results Presented on Pie Chart

3. CONCLUSION

The Analytic Hierarchy Process is a method to obtain subjective evaluations based on paired comparisons and relative judgment matrices. Although its application is fairly uncommon in the Human Factors field, the practice has found acceptance in a other fields including business, medicine, and politics. In this paper, the AHP is described as an alternative approach to statistical methods for comparing multiple design options using subjective assessments. The method is explained in part by using results from a previous study on display designs of Enhanced Vision Systems.

REFERENCES

Budescu, D.V., R. Zwick and A. Rapoport (1989). A Comparison of the Eigenvalue Method and The Geometric Mean Procedure for Ratio Scaling. *Applied Psychological Measurement,* **10** (1), 69-78.

Mitta, D.A. (1993). An Application of the Analytic Hierarchy Process: A Rank-Ordering of Computer Interfaces. *Human Factors,* **35** (1), 141-157.

Saaty, T.L. (1980). *The Analytic Hierarchy Process.* New York: McGraw-Hill Book Company.

Saaty, T.L. (1990a). How to Make a Decision: The Analytic Hierarchy Process. *European Journal of Operational Research* (Special Issue: Decision Making by the Analytic Hierarchy Process: Theory and Applications), **48** (1), 9-26.

Saaty, T.L. (1990b). Eigenvector and Logarithmic Least Squares. *European Journal of Operational Research* (Special Issue: Decision Making by the Analytic Hierarchy Process: Theory and Applications), **48** (1), 156-160.

Vargas, L.G. (1990). An Overview of the Analytic Hierarchy Process and Its Applications. *European Journal of Operational Research* (Special Issue: Decision Making by the Analytic Hierarchy Process: Theory and Applications), **48** (1), 2-8.

Vidulich, M.A. (1989). The Use of Judgment Matrices in Subjective Workload Assessment: The Subjective WORkload Dominance (SWORD) Technique, *Proceedings of the Human Factors Society 33rd Annual Meeting,* 1406-1410.

Yang, L. and R.J. Hansman (1994). A Human Performance Evaluation of Enhanced Vision Systems for Approach and Landing, *SPIE Proceedings: Sensing, Imaging, and Vision for Control and Guidance of Aerospace Vehicles,* Paper No. 2220-27.

INTENSIVE TASK ANALYSIS AND EVALUATION FOR INTERFACES DESIGN IN LARGE-SCALE SYSTEMS

ELENA AVERBUKH

University of Kassel, Laboratory for Man-Machine Systems, D-34109 Kassel, Germany

Abstract. Increasing complexity of man- machine systems and requirements to their safety, quality, enviromental compatibility and job-acceptance by all categories of users interacting with the system via man-machine interfaces lead to broad implementation of user-centered approaches to integrated automation and extensive task analysis for adequate modelling of users' behaviour, technical system functioning and of their interaction. Practical implementation of the most popular techniques, such as, e.g., hierarchical task analysis for different categories of users of large-scale systems is extremely time- and labour-consuming. This paper is focused on formalising new intensive strategies for task-analysis and -evaluation during design of adaptive interfaces for large-scale systems with several categories of users (e.g., engineers, operators etc.) . Practical example of task analysis at the biggest in Europe cement plant in Aalborg is discussed.

Key Words. Task-analyses; modelling of human-mashine interaction; large-scale systems; design of interfaces; model validation.

1. INTRODUCTION

In recent years an increasing number of research activities have been performed on advanced user interfaces design for large scale systems (LSS). Flexible interfaces which provide adaptation of system's interactive behaviour to the specific needs of human users in concrete task situations promise to improve overall systems safety and quality, as well as users job acceptance and motivation (Johannsen,1994). It, in it's turn, demands extensive task analysis for adequate modelling of human users behaviour, of functioning technical system, i.e., of application modelling and of interaction between users and system via dialogue and presentation.

Normally, the main goals of task analysis within the (re)design of constructed or operational large scale systems are the following

- identification and examining of typical task that must be and/or are performed by users when they interact with system and each other (see, e.g., Sheridan, 1992),
- evaluation of their efficiency and effectiveness for achieving concrete system goals (quality, safety, productivity, humanitarian goals etc.),
- establishing and/or modifying the task requirements,including allocation of functions between human agents and machine, as well as job organisation and interface design.

Hense, task analysis is a complicated time- and labour consuming multilevel process which covers all phases from task data collection, description and simulation to task behaviour requirements evaluation.

This paper is focused on intensive methodologies for task-analysis and -evaluation, particularly by considering cognitive models of the involved experts and thus increasing their motivation and effectiveness in general. In the next section the most traditional approach, i.e.,hierarchical task analysis and difficulties in its execution and evaluation are discussed. The third section identifies sources for intensification of this methodology. The concrete example from cement industry which particular motivated this study is given in the fourth section.

2. HIERARCHICAL TASK ANALYSIS FOR SYSTEMS INTERFACE DESIGN

One of the most popular and fundamental approaches for systems interface design is hierarchical task analysis (HTA) which is normally executed separately for each user category,e.g., engineers, operators, maintenance personnel etc.(see, e.g.,Kirwan and Ainsworth, 1992). This approach typically leads to hierarchical or mean-ends descriptions of the specific tasks using following categories (see Lind, 1981, Sundstrom, 1993, Averbukh, 1994)

goals ↬ functions ↬ tasks - means ↬ resources ↬ criteria.

These hierarchies are especially effective for interface design when they are specified symmetrically for different agents of interaction, i.e., users, human-machine interface and for technical system as well, as it is shown in the "bottle - diagram" in Fig.1,a (see Averbukh and Johannsen, 1994).

Typical straightforward execution of this methodology using interviews, walk - throughs and talk - throughs, observations and time-lines etc., has following disadvantages:

• it is time- and labour- consuming, particularly due to huge functional dimensionality and certain redundancy of hierarchical knowledge structures,

• it lacks in consideration of cognitive models of experts, which usually differ for specific categories of experts and for analysis and evaluation phases as well. It, particularly, leads to
  - non-rational use of experts available within the time and budget constraints of HTA,
  - bordom and decreasing of motivation of the experts, etc.,

• it lacks the mechanism of dynamical facilitating the needed (back and forth) transitions among levels of cognitive control and is rather uni-dimensional and sequence-oriented. That
  - provides only a "static" framework of the job composing tasks and to a large extent ignores, e.g., the need for and effects of the forgoing "intention formation" for successful job performance,
  - is somehow effective for traditional hierarchical job organisation, but not for modern tendencies towards flexible team organisation structures (see, e.g., Meshkati, 1991).

Considering these obstacles, the next section is focused on several specific strategies for intensification of the HTA and evaluation of its results.

3. INTENSIVE STRATEGIES FOR TASK ANALYSIS AND -EVALUATION

3.1. Reduction of Complexity and Flexibility - "Chain Diagrams"

As it is shown in Fig.1,a, means-ends hierarchies can be schematically associated with the following questions for task- analysis and - evaluation:

goals♦<u>why</u> ↬ functions♦<u>how</u> ↬ needs <tasks ♦ <u>when</u> - means(♦ <u>what</u>)>

Further, for the sake of simplicity we shall conditionally associate means (informational, communicative, technical etc.) and tasks on both activities- and specific "key-stroke" actions-levels with the only question <u>when</u> (to do what) related to generic *needs*.

These questions give a frame for analysis of so called *procedural* task-knowledge or task - characteristics. Analysis of available resources for execution the particular task and of related criteria leads to specification of so-called *situational* task characteristics (Averbukh, 1994). All together they allow to restructure fundamental and redundant hierarchical "bottle-diagrams" into flexible "chain-diagrams", taking into account different end-users or other agents of interaction (question <u>who</u>), as it is shown in the Fig.1,b.

Chain task-knowledge representations give new opportunities for coping with the problems specified in the section 2, as it is shown bellow.

3.2. Rational "Short-cut" Strategies

Rationality and effectiveness of execution of task analysis can be significantly improved by considering typical "short-cut" reasoning strategies for different groups or categories of users, e.g., of supervisory and control systems (see, e.g., Rasmussen, 1981,1986). The examples of such strategies for two user groups, i.e., operators of supervisory and control systems and engineers during HTA for normal and emergency situations are schematically depicted in Fig.2,a,b (so-called Z - and δ – strategies).

The task analysis experience (see, e.g., Heuer et al,1993)shows quite strong "zigzaging" preferences, i.e., Z-strategies for both categories of users in analysing situational vs procedural task characteristics within normal situations' scenario (Fig.2,a). In trouble shooting both user categories prefer knowledge-based δ – strategies of initial task-situation assessment (resources and criteria) with following selection of appropriate procedural task-characteristics (how-when) and "double-checking" of criteria (Fig.2,b). For cooperative trouble-shooting scenario one of the possible short-cut analysing strategies is depicted in Fig.2,c as an ε– strategy.

Thus, preliminary structurisation of interaction knowledge using "chain - diagrams" and *user-, scenario- and HTA phase - oriented* short-cut strategies, either approven, or hypotetical will allow knowledge engineers to intensify HTA significantly by following adequate to the involved experts reasoning strategies and avoiding boring redundancy and "unusual" short-cuts, e.g., in interviews etc. It also simplifies multi-dimensional task-analysis for team problem-solving scenaria.

In the next paragraph the way to intensify evaluation by executing rapid "in-process inspection" of task knowledge is discussed.

3.3. Cross-Validation of Task-Knowledge between Different User-Groups

The real-life experimental conditions for execution and especially for evaluation of the results of Hierarchical Task Analysis in systems interface design are usually extremely restricted. These constraints originate not only from the limited budget and strictly planned project life-cycle, but also from restricted availability of experts, impossibility (e.g., for engineers or managers) to isolate themselves from daily job problems and to concentrate completely on task analysis, high influence of personal psychological factors, lack of formalisation and of fixed planning in HTA procedures due to different kinds of uncertainties and unforseen factors etc.

This leads to decreasing of quality and reliability of the acquisited and formalised task knowledge. Bellow possible methodological solutions for improving total quality of both analysis and evaluation processes are specified.

• Preliminary identification (prediction) of locally stable cognitive strategies and preferences in communication styles during both task-analysis and -evaluation phases for
  - specific user groups,
  - particular experts available within particular HTA phase.

Different test-procedures supported by modern knowledge-based tools for knowledge acquisition (see overview, e.g., in Gavrilova and Tschervinskaya, 1992) and specified above chain-representations can be used for this purpose.

• Adaptation of HTA schedule to these identified possibilities/preferencies. E.g., neighbouring categories of users within traditional job hierarchies usually have knowledge about the same task characteristics with different level of abstraction and uncertainties and also, possess and/or prefer "opposite", i.e., goal- and data-driven reasoning strategies. It should be appropriately considered in planning and scheduling of both analysis and evaluation phases.

• Planning of the possible switching points within the analysis schedule for rapid "in-process" inspection and/or completing of the acquisited knowledge by means of so-called cross-validation. It can be also called by analogy, concurrent engineering approach to HTA (Hsu et al, 1994).

E.g.,experts from the predecessor user-group within traditional job hierarchy in LSS often prefer to specify typical (rough) situational and task characteristics for the successor user-group. This together with some observations of performance of the users-successors can be used by knowledge-engineers for preliminary "chain" knowledge representation with the following validation and/or completing by the experts from the successor user-group. The given bellow example has proved this hypothesis.

Mentioned above local cognitive models for different user groups as well, as cross-validation approaches are schematically shown in Fig.3,4.

4. EXAMPLE FROM CEMENT INDUSTRY

The specified above methodological principles were partially implemented and tested during rather intensive task analysis. It was executed by the team of three knowledge engineers from the University of Kassel together with participants from FLS/A at the Aalborg Portland, one of the largest cement plants in Europe (see Heuer et al, 1993).

It was conducted in the frame of the BRITE/EURAM AMICA Project during 3.5 months for two user groups, namely control room operators and engineers (both control room- and commissioning engineers).

First unstructured interviews and getting acquainted with the plant and control rooms were organised mainly as "group" interviews (2 days) with two operators, and three engineers. Later during the structured interviews and cooperative scenario construction the 3 sessions with 2 engineers (3 hours each) and sessions with 2 operators (2 whole day sessions with one operator from 6 to 8 hours) and 5 sessions with another lasted in average 2-3 hours were carried out.

The unpredicted bordom was observed during discussion of concrete procedural details (activities-actions) with operators for specific scenario "Start-up the mill with OPC cement". Within very restricted schedule engineers have shown significant interest in analysing and verifying tasks of operators (but not on the "key-stroke" level).

Rather intensive execution of this task analysis, as well as parallel activities within the AMICA project at Santa Gilla power plant by the knowledge-engineers of CISE and ENEL were highly respected by industrial representatives. These activities motivated and in general approved the efficiency of the above methodological guidelines.

5. CONCLUSIONS

The problems discussed in this paper have extremely high importance for large scale systems design in achieving overall quality, safety, cost reduction, as well as users satisfaction during all design phases. Further research is needed in formalising concurrent engineering approach to task-analysis and evaluation and in developing appropriate knowledge-based tools and support technologies. Preliminary Expert Modelling, particularly using flexible "chain" knowledge representations for intensive execution of task analysis with rapid validation seems to be indispensible for systems´ interface design, especially within new paradigms of team job organisation.

6. REFERENCES

Averbukh, E.A. Task-Orientation of Human-Machine Interaction in Distributed Systems, Proceedings of 3rd IEEE International Workshop on Robot and Human Communication, Nagoya, Japan, July 18-20, 1994 a.

Averbukh, E.A. Multidisciplinary Task Knowledge Engineering, Proceedings IFAC Conference on Integrated Systems Engineering, Baden-Baden, Germany, September 27-29, 1994b.

Averbukh, E.A. and G. Johannsen: Intelligent Human-Machine Communication and Control for Industrial Processes, Proceedings of First Asian Control Conference, Tokyo, July 30 - August 2, 1994.

Gavrilova, T.A. and K.R. Tschervinskaya, Knowledge Elicitation and Structurisation for Expert Systems, 1992, Moscow,Radio & Communication (in Russian).

Hsu, C., Gerhardt, L., Spooner, D. and A. Rubenstein, Adaptive Integrated Manufacturing Enterprices: Information Technology for the Next Decade, IEEE Transactions on SMC, 1994, V.24, No.5, pp. 828-837.

Heuer, J., S. Borndorff-Eccarious, E.A. Averbukh Task Analysis for Application B, Internal Report IR1-04-BRITE/EURAM Project 6126 (Nov.1993).

Johannsen, G. Design of Intelligent Human-Machine Interface, In Proc. of 3rd IEEE International Workshop on Robot and Human Communication (RO/Man'94), Nagoya, July 1994.

Kirwan, B. and Ainsworth, L.K. A guide to task analysis, Tailor & Francis, 1992.

Lind, M. The Use of Flow Models for Automated Plant Diagnosis, in Human Detection and Diagnosis of System Failures, J. Rasmussen and W.B.Rouse, eds. London, Plenum Press, 1981, p.411-432.

Meshkati, N. Integration of Workstation, Job, and Team Structure Design in Complex Human-Mashine Systems: A Framework, International Journal of Industrial Ergonomics, 1991, 7, pp. 111-122.

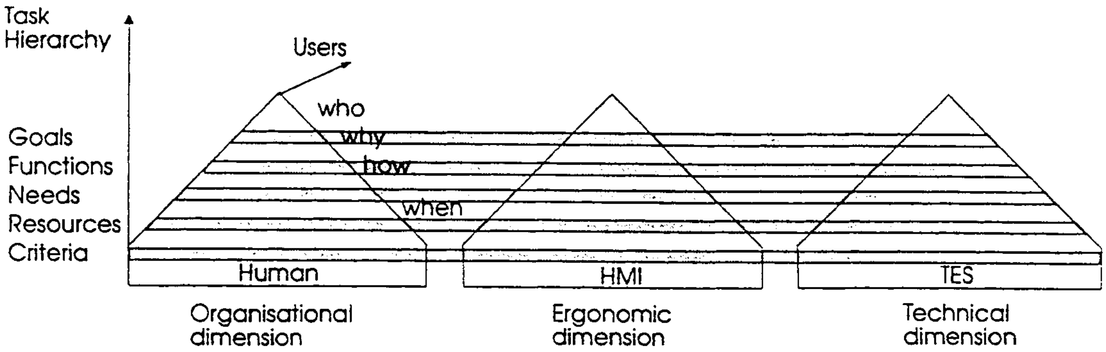
Rasmussen, J. Models of mental strategies in process plant diagnosis. In: Human Detection and Diagnosis of System Failures, edited by J.Raasmussen and W.B.Rouse, 1981, New Jork:Plenum Press.

Rasmussen, J. Information Processing and Human-Mashine Interaction; An Approach to Cognitive Engineering, 1986, Amsterdam, Elsevier.

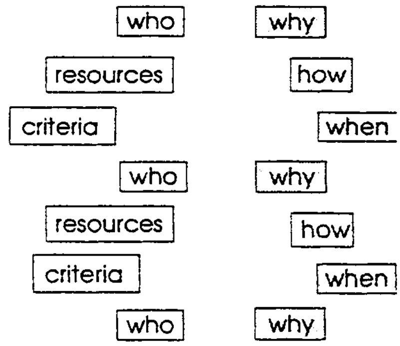
Sheridan, T. Telerobotics, Automation, and Human Sypervisory Control. Cambridge,MA:MIT Press (1992).

Sundstrom, G.A. Model-Based User Support: Design Principles and an Example, In Proc. IEEE/SMC'93, Le Touquet, October 1993, p.355-360

The presented work is partly supported by the Commission of the European Union under the BRITE/EURAM II programme in Project 6126 (AMICA: Advanced Man-Machine Interfaces for Process Control Applications). The consortium partners are: CISE (Italy), ENEL (Italy), FLS Automation A/S (Denmark), University of Kassel (Germany), Marconi Simulation and Training (Scotland).



a) "Bottle"-diagram



b) "Chain"-diagram

Fig. 1 Integrated vs distributed task knowledge structurisation

- Procedural task characteristics
- Situational task characteristics

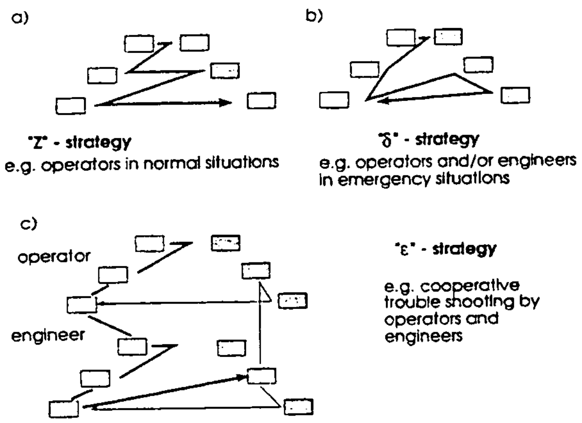


Fig. 2 Typical "shortcut" strategies in task analysis

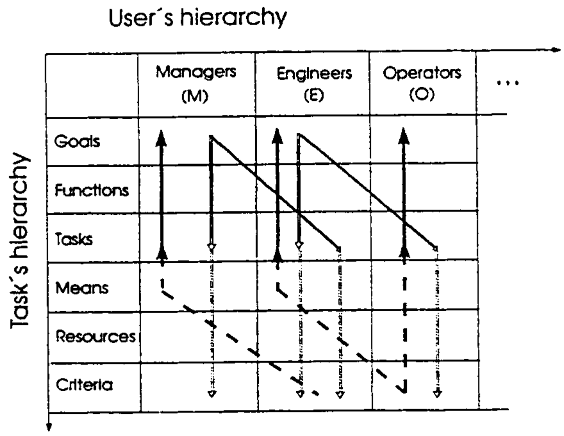


Fig. 3. Locally "stationary" cognitive models of systems´users

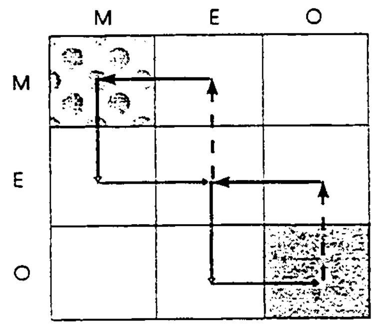


Fig. 4. Strategy for rapid cross-validation of task-knowledge

- goal-driven analysis
- data-driven analysis
- relates to low levels of Task´s Hierarchy: Means-Resources-Criteria
- relates to high levels of Task´s Hierarchy: Goals-Functions

MEDIATION OF MENTAL MODELS IN PROCESS CONTROL THROUGH A HYPERMEDIA MAN-MACHINE INTERFACE

J. Heuer, S. Ali and M. Hollender

University of Kassel, Laboratory for Human-Machine Systems, D-34109 Kassel, Germany

Abstract: In this paper a new paradigm is described which enhances conventional Supervisory and Control (S&C) systems to support- and learning systems. Those hypermedia S&C systems provide the operators with multiple views of different aspects of the process and can be used as a corporate knowledge base, where operators are encouraged to add their experience and to retrieve the knowledge of others. The concept fosters a continuous learning process during every day work.. The structure of the hypermedia S&C system is strongly influenced by the causal interconnections between the process variables, helping operators to develop a correct mental model of the process.

Key Words: Mental models; hypermedia; process control; man-machine interface; learning.

1. INTRODUCTION

In recent years there has been a tendency in process control towards increased complexity, both in the connections of components and in the overall process automation schemes. This situation highly increases the demands on the operators for effectively running a given process. Therefore the quality standards in process control do not depend solely on the characteristics of the technical process, but rely heavily on the educational standards of the operators (Sage, 1992). Additional efforts are needed to provide the knowledge and the skills that are necessary to run a complex process, especially in critical situations (Brehmer, 1987).

In this paper a framework will be proposed that covers both aspects: the increased quality of the man-machine interface and the process visualisation, as well as the mediation of knowledge. The concept of mental models (Johnson-Laird, 1983) as the representational form of knowledge will be used and it is assumed that the existence of correct models is a prerequisite for successful process control (Johannsen, 1993). The process visualisation will not merely be used as a "window to the process", but will serve as a support- and learning system, which provides the operators with multiple views showing the different aspects of the system. To achieve this goal, a concept of the process visualisation system as a hypermedia-system will be proposed. This system will support the operators on different cognitive levels during supervision and control and will also serve as a source for corporate knowledge (Steels, 1993) that is constantly enhanced. The proposed concept fosters a continuous learning process during everyday work, which makes the learning itself highly effective (Lave & Wenger, 1991). Increasing the qualification and capabilities of the operators is one way to reduce what Bainbridge (1987) calls "the ironies of automation". Although this provides no safeguard against "latent errors" (Reason, 1990) it does help reducing the active operator errors. Another major goal is to improve the motivation or the complete and full engagement of the person in pursuit of the end cause of the activity (Laurel, 1986), a notion for which work psychology has coined the term *task orientation.*

2. LEARNING OF COMPLEX SYSTEMS: A DESTILLATION COLUMN

Large technical processes represent highly complex, dynamic systems. Operators working under such conditions are expected to have stable and correct models of the interconnections between the process variables and to act accordingly. An extensive technical education as well as an intensive training with the technical process are necessary in order to develop what is called expertise (Johannsen, 1993).

To model a realistic process under laboratory conditions it is necessary to use high-fidelity simulations. One such simulation was developed at the University of Stuttgart (Gilles, Holl, Marquardt, Schneider, Mahler, Brinkmann & Will, 1990). It models a chemical destillation column for the separation of benzene and toluene. A schematic flow diagram can be found in Figure 1

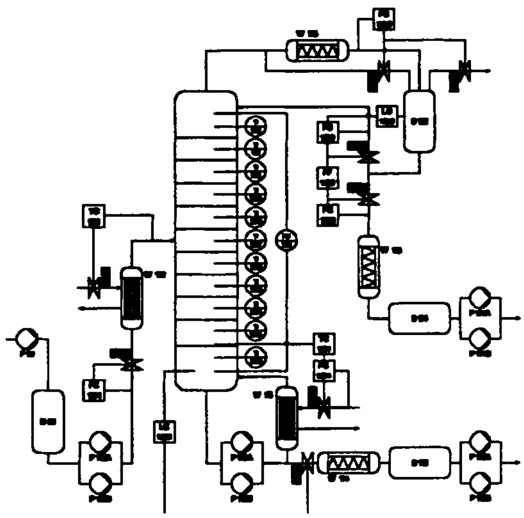


Figure 1: Flow diagram of a destillation column

3. MENTAL MODELS AS A PREREQUISITE FOR PROCESS CONTROL

The quantity and quality of mental representations an operator holds about a certain process form the basis of her successful acting even in difficult situations. These representations come in different varieties, namely as analogue representations, procedural *schemata* and *scripts* and declarative *networks* (Anderson, 1980). The different terms and notions of the representations will be subsumed under the term *mental model*. These models have both predictive and explanatory power (Norman, 1983), e.g. an operator act on a process by recognising the actual process situation and then choosing the appropriate mental model to act upon. These models (in this case schemata or scripts; Schank and Abelson, 1977) contain information that allows proper identification of the right schema for a given situation (i.e. a conditional part) and also fixed sets of actions that can be easily executed when the conditional part is matched. These schemata can be hierarchically composed of different sub-schemata, depending on the complexity of the situation. Hacker (1986) states that the correctness and level of discrimination of these models determines the quality of the operators actions, and that the quality of acting in the work context is mostly dependent from adequate underlying models.

Different studies in the area of problem solving in dynamic systems have shown (e.g. Dörner, 1986, Brehmer, 1987; Funke, 1992) that human problem solvers have great difficulties in learning such systems by active manipulation and in building an adequate mental model. This seems to be especially the case with causal models, e.g. models in which the interconnections between variables in a system are represented. According to Funke (1992), these models develop in three successive steps. The first steps involves the building of general hypotheses by the problem solver. These hypotheses only contain assumptions about general connections between certain variables. The second, semi-qualitative step, makes assumptions about the direction of the established connections. Finally, on the highest, quantitative level, problem solvers exactly evaluate the strength of these connections. The ability to build these models deteriorates if the system gets more complex and dynamic. Yet, exactly this happens in industrial and especially chemical processes and systems. For this reason it is necessary to support the operator in identifying system connections by using the adequate representation techniques.

While the knowledge about the underlying structure of the system variables is necessary e.g. for trouble shooting and other problem solving activities, there also has to be support concerning the control-ability of the operators. As Broadbent, Fitzgerald and Broadbent (1986) have shown, there can be a considerable difference between the system knowledge somebody is able to verbalise and the quality of her control actions. This means that somebody who is able to correctly control a dynamic system does not need to have an explicit causal (mental) model and vice versa. Broadbent *et al* (1986) explain this by proposing two different problem solving strategies: The strategy of *model manipulation* assumes that the problem solver possesses a model about the system which matches the "real" situation as closely as possible. All actions and explanations for system states can be derived directly from this model. Using the strategy of *situation matching*, the problem solver has a repertoire of correct actions which are linked to certain situations. This enables the problem solver to show a good performance when confronted with most of the common system states, but she is not able to perform predictions or reasoning concerning new situation on this basis. This strategy will fail in situations not yet available in the operator's repertoire (e.g. rather seldom system failures in chemical plants), and may result in different sorts of errors as described by Reason (1990).

Therefore the main goal of educating process control operators has to be the mediation of correct and stable mental models, on which both actions and problem solving activities can be performed. The knowledge to be taught has to contain knowledge about the technical process itself, resulting in a network structure or causal model. It is also necessary to provide knowledge about "how-to" control the process in many different system states.

In this paper, a MMI will be presented which incorporates the presentation of different types of mental models and thus serves to mediate knowledge on the different levels proposed. The following parts of the MMI will be presented in detail:

- a dynamic causal model
- a qualitative simulation enabling the operator to explore the process in different situations
- a dynamic blackboard- and help system.

4. DIFFERENT MODELS OF THE PROCESS

The contents which are represented in the MMI are essentially based on different Mental Models of the user concerning aspects of the process. The information which forms the basis for the representations was gathered by studies of written material, questioning experienced operators and an extensive task analysis. This gathering of content to be presented forms one part of a circular process, the goal of which consists in a tight match of the *design model*, the *User Model* and the *System Image* (Norman, 1983). Figure 2 gives an overview of the process and the different models involved.

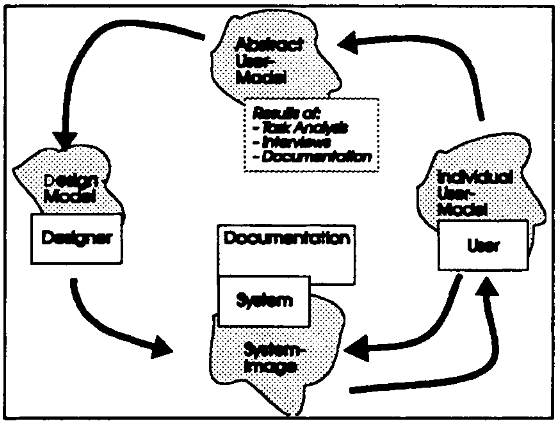


Figure 2: Different Models of a process control system (after Norman, 1986)

The data stemming from the analysis is necessarily incomplete and does not allow for an exact match of the *individual* user model. Therefore it is important to choose the level of resolution (Shen and Leitch, 1992) for building the *abstract* "typical" *user model* (Norman, 1983). Level of resolution in this case refers to the absolute number of existing variables in the technical system (about 120 in the destillation column) and the number of salient variables (the cognitively most prominent variables). The model that was identified in this way contained merely the six input and six state variables that account for most of the dynamics in the process.

The *design model* represents the conceptual model of the system to be designed. It is the task of the designer to translate these abstract models into the *system image*. Ideally, the system image contains different graphical representations of the original user models. The presentation of these models should provide information of as many different aspects of the technical system as possible.

5. USING HYPERMEDIA FOR MAN-MACHINE INTERFACES IN PROCESS CONTROL

Many processes have hundreds or even thousands of variables. S&C systems must group information about these variables into screens and relate the screens with each other. This is an amazing analogy to the nodes and links of the hypermedia world. Indeed, modern S&C systems can be more than just a mimicry of the former switchboards. Today, screens can be enriched with knowledge. They can be used as a means for communication between different shifts or between experienced and junior operators. A well defined structure offers "predictable paths to expected information" (Herrstrom and Massey, 1989). This structure should be essentially hierarchic but can be enhanced with other task-oriented relations like causal relations (Hollender, 1994).

6. HYPERMEDIA - BASED INTERFACES AS LEARNING ENVIRONMENTS

The use of hypermedia-based man-machine interfaces serves mainly one purpose: To provide the operator with different views of the same process. The reason to do so comes from the special abilities and weaknesses of human information processing. Anderson (1983) applies different mechanisms on the choice of a specific production rule. One of the most important of these mechanisms is the degree of match between the given situation and the conditional term of a production rule. The more specific and exact this match is, the more likely is a production to fire. Under the assumption that the choice of situation-specific schemata for controlling actions is based on similar mechanisms it can be said that specific schemata become activated when their "slots" match the data in the given situation. If no schemata exist that have a "perfect" match, decision procedures are involved that may possibly lead to the choice of the wrong schema (e.g. "frequency gambling" or "similarity matching", Reason, 1990). The more differentiated and specific the existing schemata are, the more precise the matching procedures can be. The hypermedia-based MMI provides the mechanisms to tune the existing schemata into a more specific form. Also, the different views of the system can be used to provide more and different information for the matching procedure in the case of uncertainty (i.e. no conditional parts are exactly matching).

If, for example, an operator encounters a situation where the system for some reason gets unstable, the first step would be to recognize the *cause* for the instability. This would be the fault hypotheses of the operator. Using the standard (topological) display, the only data for accepting or rejecting the operators fault hypothesis comes from two sources: The display and the operators knowledge. Since these sources are not independent, the operator will selectively search the display for confirming evidence. This results in a fair probability of erroneously interpreting the only existing data source (the display). On the other hand, in using the multimodal (multimedia) displays, the operator may gather his data on a *causal*, a *procedural* and a *goal oriented* basis. In addition to this (always under the assumption that there is enough time available) she might also request data from the blackboard- and help system. This cooperation of different perspectives differentiates the data the operator is using for confirming her hypotheses, thus leading to a much firmer basis for the following course of action to take (i.e. accepting a given hypotheses as the conditional part of a production or schema).

6.1. The Causal Model and Simulation

The "backbone" of the proposed hypermedia-network is a causal model of the system. Although the often used topological representation of a process, providing a schematic "map" of a system, proves to be very useful in directing field operators and providing a "mental map" for the operators, this presentation is not sufficient. One of the major drawbacks of the topological representation is the fact that variables (or components of the system) that *functionally* have no or very little influence on each other may be shown closely grouped together; on the other hand there may be very tightly coupled variables which are presented far apart from each other. To overcome these difficulties, the causal model was developed in addition to the topological presentation.

This model consists of different nodes, each of which is representing a specific system variable. The nodes are distinguished in two different types. The first type - direct nodes - represents a process variable that can be directly manipulated by the user. The second type - indirect nodes - stands for variables that can only be influenced via the direct ones. The links between the different nodes have a double functionality. On the one hand they represent the *strength* of influence a node is exerting on another node. On the other hand, links are used to represent the time it takes for a manipulation on one node to influence the connected node. The visualisation of the network enables the operators to set the visual focus to the node (variable) they are currently interested in. This node is presented in the centre of the screen. The nodes exerting influence on the central node are presented to the left, ordered according to the strength of their influence and the response time. Variables being influenced by the central node are shown on the right, ordered accordingly. The operators can select every preceding or succeeding node, thus being able to follow complete cause-effect chains. The visualisation of the causal model is intended to support the supervisory *planning* step. Figure 3 presents an overview of the qualitative causal model.

6.2. The Hypermedia Help- and Blackboard System

In addition to the multimodal presentation the operators will be supported by a dynamic blackboard- and help system. The information contained in this system may contain textual, pictorial and dynamic elements. Every operator can add new elements to the blackboard which are immediately available to all other personnel. Information contained in this blackboard will be evaluated in regular sessions and will be eventually become integrated in the help-system network as new nodes and links.

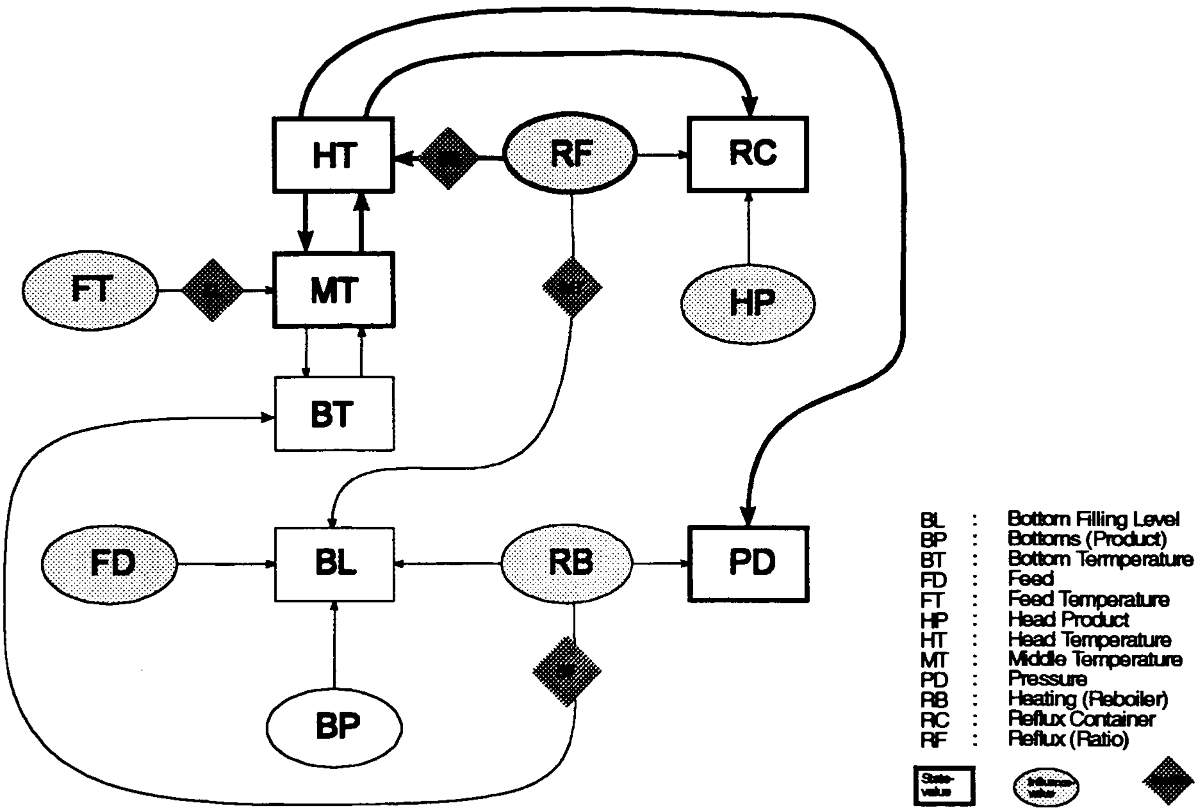


Figure 3: Qualitative causal model of the destillation column

7. IMPLEMENTATION CONCEPT OF THE INFO SYSTEM

The *INFO-MS (Information, Navigation, Failure handling and Organisation - Management System)* is implemented in a shell structure. The basic elements of the kernel are module- and process variable objects. The element organisation is managed with the help of dynamically linked lists. The initial data for the dynamic lists can be loaded from an ASCII file (process model file) or from an object-oriented database. The process model file contains the structure of the plant in a hierarchical form.

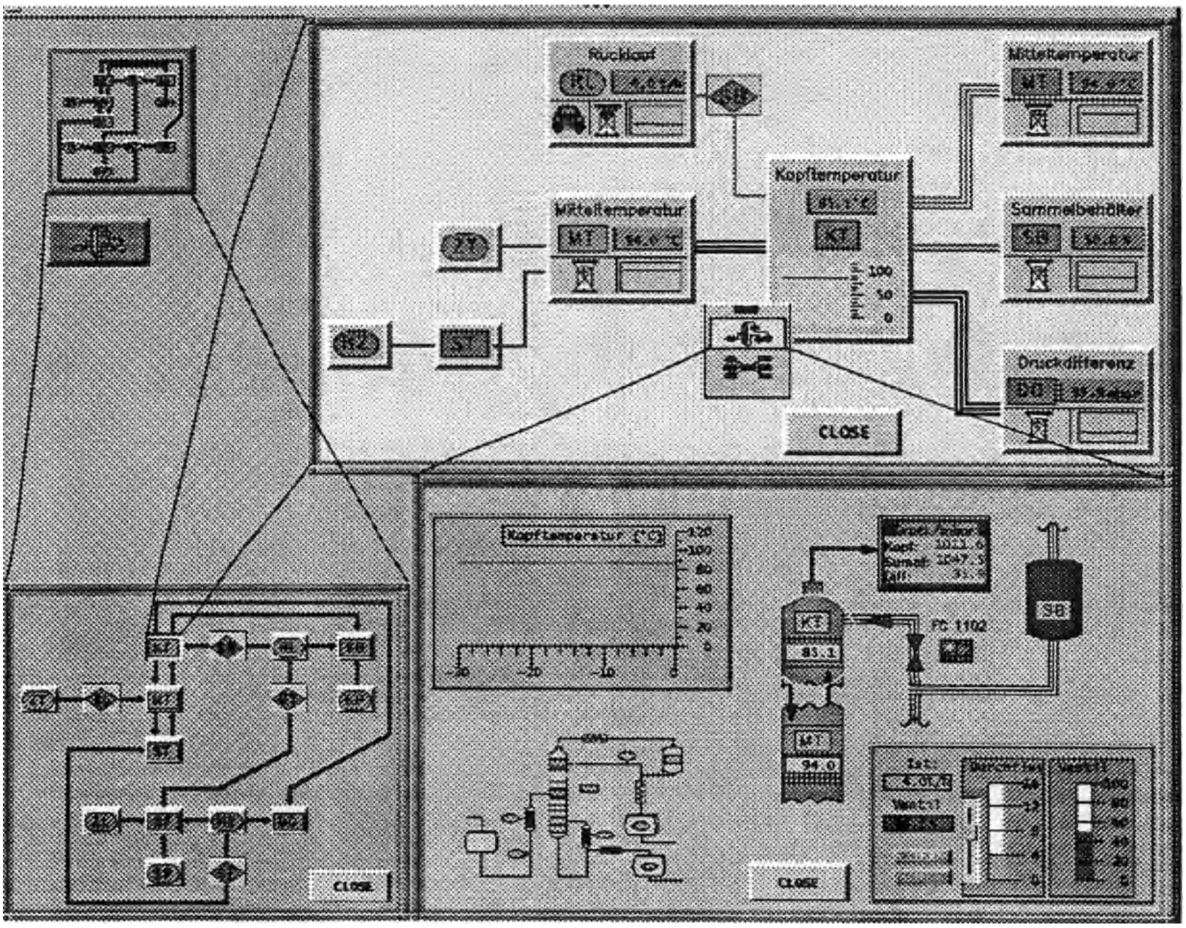
The *INFO* MMI visualises the process in several representations, such as topology, causal nets or goal oriented models, and gets the animation data from the *INFO-MS*. The *INFO-MS* gets the actualisation data either from the technical process, animation objects (simple simulations) or the hypermedia system *HypMed*. The animation data of *HypMed* were recorded previously from critical process situations and managed from a play-back driver. Thus, there is a simulation database of critical process situations that can be used any time for training purposes. The structure of the database of the hypermedia system *HypMed* is based on the qualitative causal model, process states and goal oriented hierarchies. The qualitative model contains qualitative objects such as process variables and connections. Every qualitative object contains different attributes such as a fuzzy variable, fuzzy sets, a local fuzzy rule base about change-of-rates, a delay time etc. The values off attributes are suitable to animate the causal-oriented user displays.

Figure 4 shows a minimised causal model (lower left), the expanded links display (upper right) and the corresponding topological view (lower right) is An example of the INFO - MMI.

8. CONCLUSION

In this paper it was shown how hypermedia-based man-machine interfaces can be used to create a manifold, yet consistent representation of a complex technical process. Through the use of multimodal presentation techniques it is possible to mediate mental models concerning different areas of knowledge, namely procedural and declarative mental models. The mediation of these different models provides the process control operators with a solid based for their supervisory and control tasks (Sheridan, 1993).

Currently, extensive experiments are carried out to determine the amount and types of knowledge mediated by the interface. This is done by using different learning environments. The effect of these different representations will be assessed with help of several relevant scenarios. The experiences



already made with the new representation techniques give hope to a successful application of the complete system in the future.

REFERENCES

Bainbridge, L. (1987).The ironies of automation. In: J. Rasmussen, K. Duncan and J. Leplat: *New Technology and Human Error, 111-120.* Chichester: Wiley.

Brehmer, B. (1987). Development of mental models for decision in technological systems. In: J. Rasmussen, K. Duncan and J. Leplat: *New Technology and Human Error, 111-120.* Chichester: Wiley.

Gilles, E.D., Holl, P., Marquardt, W., Schneider, H., Mahler, R. , Brinkmann, K. & Will, K.-H. (1990). Ein Trainingssimulator zur Ausbildung von Betriebspersonal in der Chemischen Industrie. *Automatisierungstechnische Praxis, 7,* 344-350.

Herrstrom, D. S., Massey, D. G. (1989) Hypertext in context. In Barrett, E., *The Society of Text.* MIT Press, Cambridge. pp. 45-58.

Hollender, M. (1994). *Diagnosing faults co-operatively and in context.* Preprints SAFEPROCESS'94. Vol. II, pp. 643-648

Johannsen, G. (1993) *Mensch-Maschine-Systeme.* Berlin: Springer.

Johnson-Laird, (1983). *Mental Models.* Cambridge, MA: Harvard University Press.

Laurel, B. K. (1986). Interfaces as Mimesis. In: D. A. Norman and S. W. Draper: *User Centered System Design.* Hillsdale, NJ.: Lawrence Erlbaum.

Lave, J., Wenger, E.(1991). *Situated Learning: Legitimate Peripheral Participation.* Cambridge; MA: University Press.

Perrow, C. (1984). *Normal Accidents: Living with high-risk technologies.* New York: Basic Books

Reason, J. (1990). *Human Error.* New York: Cambridge University Press.

Sage, A. P. (1992). *Systems Engineering.* New York: Wiley.

Schank, R.C. and Abelson, R.P. (1977). *Scripts, plans, goals, and understanding.* Hillsdale, NJ: Earlbaum Associates.

Sheridan, T. (1993). *Telerobotics, Automation and Supervisory Control.* Cambridge, MA: MIT Press.

Steels, L. (1993). *Corporate Knowledge Management.* AI-Memo, Vrije Universiteit Brussel.

REAL TIME EXPERT SYSTEM IN PROCESS CONTROL : INFLUENCE OF PRIMARY DESIGN CHOICES.

V. Grosjean

Ergonomics and Industrial Psychology Department
National Institute for Research and Safety (I.N.R.S.)
B.P. 27 Av. de Bourgogne
54501 Vandoeuvre F.
tel.: +33. 83.50.20.00.
e-mail : grosjean@inrs.fr

Abstract : This paper presents a comparison of two supervisory control tools based on real-time expert system technologies. The two tools are used in the same kind of environments, but have been developed using contrasting design philosophies. The first one has the objectives of fitting the technical constraints of the process and of being easily adaptable to all similar processes while the second has the objective of developing a maximal consistency between the process control strategies of the tools and the local process control strategies of the operators. The first philosophy offers better performance in terms of both acceptation and use. Explanations and consequences for generalisation are discussed.

Key-words : system design, expert system, decision support system, knowledge engineering, process control, supervisory control, consistency.

1. INTRODUCTION

This paper is focused on the problem of the use of real-time expert systems in process control environments. It is based on observations gathered on two industrial applications of expert systems.

Our interest in these observations is connected to a theoretical problem, namely the problem of the consistency between the user - in our situation, the operator - and the Expert System used in the task at hand.

The Human Factors' point of view of this problem of consistency between intelligent support systems and human operator was quite clear until some years ago ; the general recommendation for the designers of such systems was twofold :

- on the one hand, try to maximise the correspondence between the human-operator's problem solving and that of the expert system. That is to say that the expert system should control the process in the same manner, and by applying the same rules, as the operator does ;
- on the other hand, adapt the presentation of information by the interface in such a way that it conforms to the operator's expectancies.

This point of view is central to Human-Centred Automation, and corresponds to the now classical User-Centred Design approach used in HCI (Norman & Drapper 1986). It has been an important point in promoting the evolution of system design from one centred exclusively on technical aspects to one taking the human operator into consideration. One of the main advantages attributed to consistent tools is that they improve the user's trust in the machine (Muir 1987), in so far as the operator is supposed to prefer tools that are adapted to his way of working.

But now, some Human Factors researchers have reconsidered this position, starting from different considerations :

- from a theoretical point of view it has been argued (Moray 1987) that an expert system that tries to reproduce human cognitive characteristics could be dangerous in some situations. For instance, such a system could reinforce the rigidity of problem-solving strategies ("cognitive lock up") of operators in diagnostic activities :

"When a fault occurs, operators do not usually explore the environment and update their mental model. Instead of being triggered by discrepancies to pursue induction, they desperately try to fit data to their existing models, and are typically inflexible.[...] That is why intelligent decision aids, which do not become stuck, are required. It is also why expert systems which mimic the properties of humans, are exactly what are not required as intelligent decision aids".(Moray 1987). To avoid such fixation errors (Woods 1986, De Keyser et Woods 1989) it is then better to design expert systems that use strategies that are different of those used by human operators.

- from the experimental point of view, some data (Lehner 1987) also suggest that the particularities of real time expert systems lead to design choices different from those available for classical off-line expert systems. Lehner sees three reasons for this : 1) the temporal constraints of expert systems that impose (on the operator) a rapid understanding of its inference ; 2) the inter-individual differences between operators in so far as their expertise is concerned, leading to differences in the interpretation of the expert system decisions ; 3) the fact that the operator does not input the data in the system and that he could not know their values. For Lehner, the greater the distance between the operator's reasoning modes and those of the system, the better the result of their co-operation. Meanwhile, he insisted on the necessity of acquiring a good conceptual understanding of the expert system's principles and reasoning modes.

Nevertheless, empirical data related to the use of real time expert systems in real situations are very difficult to find. Information related to the effects of the basic design choices on the efficiency and acceptation of an application in a factory come mainly from experimental applications (Mitchell and Miller 1983, Minault and Penalva 1993) and are hard to transfer to field applications. The primary objective of the present work is to give some information concerning the effects of these primary choices on performance, and on both individual and collective acceptance of real-time expert systems.

2. CONTEXT OF THE RESEARCH

A comparison of design and use of two real time expert system applications will be presented here. These applications were developed in two French cement groups (surprisingly, some of the first and most advanced applications of supervisory control tools based on real time expert systems are running in the cement industry).

This comparison was carried out during the course of a larger survey conducted in France on the application of process control tools, and ergonomic problems in supervisory control situations (Grosjean 1994a et b).

The first part of the study was conducted with the technical support team of the two cement groups. The second part was carried out in one of the factories of the group where the expert system was used. The point of view of the people in charge of the conception of the system, the knowledge elicitation, and the construction of the Knowlede Base was considered, as well as that of the users (final users and engineers responsible for the updating of the tools).

We have tried to understand the consequences of the design choices in the two groups from a systemic perspective, that is to say in their cognitive, motivational (acceptation of / trust in the technology and the tool), and organisational standpoints.

3. METHOD

The first part of this investigation was conducted through interviews in the technical support team that is in charge of the development of the expert system platform and of the implementation of the expert systems in factories. The approach was semi open-ended. The following points were investigated : philosophy of the supervision praised by the design team, characteristics of the presentation of information, ergonomic recommendations adopted, ...

The second part of the investigation was carried out in one factory of each industrial group through interviews and observations :

- Observations and interviews with plant managers, aimed at identifying the process and the expert system (nature, complexity, level of coupling, stability of the process, nature and importance of control automation, history of the implantation, ...).
- Observations of the expert system and supervisory software (information coding, plan segmentation, kind of automation used, complementary tools used, information and commands completeness, ...).
- Interviews with people involved in the updating and use of the expert system.
- Semi open-ended observation of the supervisory activities, with verbalisation (to identify the comprehension of the information and of the behaviour of the tool, to identify the level of control used, the level

of proceduralisation in the task, the nature of the activities, ...).

- Observation of visible ergonomic problems (vigilance, over informational load, diagnosis difficulties, ...).
- Interviews with operators guided by the problems encountered (small incidents) during the supervisory control activity.

4. RESULTS : COMPARISON OF THE TWO SITUATIONS

The results presented here are only an excerpt of the set of results concerning the two applications. It should be noted that the two processes are very similar in scope and in complexity.

4.1. Description of the functions of the expert systems

In the two cases, the expert system is connected on-line to the process and to the low level automatons, and the task-specific data are thus automatically gathered by the expert system. The system produces its inference with a short cycle, namely about every 3 minutes.

The expert system could run in "open-loop" or in "closed-loop" mode. When running in open-loop mode, the results of the inferential process (evaluation of the state of the process, action requirement, ...) are provided to the operator who could use it or not to make his decisions. When running in closed-loop mode, the results of the expertise are used by the expert system to control the process without the intervention of the operator, but the operator is still informed about the diagnosis and decision taken by the system.

Thus, in open-loop mode the system is running as an intelligent assistant, and in closed loop mode it is running as a semi-independant supervisor *and* as an intelligent assistant.

4.2. Information provided

The two interfaces present a number of similarities in so far as the nature and the structure of information are concerned. Nevertheless, the first interface makes a greater use of graphical coding than the second. For both supervisory control applications (including the expert system + classical supervisory tools), three levels of information could be distinguished, each level corresponding to a specific scope of activities :

- The first level, corresponding to the emerging information built by the E.S. Information at this level is very condensed and very global. It relates for instance to the state of the three successive stages of the process. This structure information is used by the operator to verify that everything is normal in the three main sections of the process, or to identify the tendencies of evolution of these phases.
- At a second level, information corresponds to traditional automation applications. This information concerns the state and the evolution of the basic parameters of the process and of the automation tools. That is exactly what can be found in process control situations where no expert system is used. The operator will use this information when a problem arises, when he thinks that a problem could arise, or when he thinks that the information presented at the first level is not completely reliable (and must be cross-checked !).
- At a third level, information corresponds to the inferential processes carried out by the expert system. This information is presented as inferences trees that can be at different distances from basic process (and automaton) information. They are used by the operator when he thinks that something is going wrong in the inference process carried out by the expert system.

In the first firm, the three levels of information correspond to three different VDUs. In the second one, the first and the third level are provided by the same VDU and the second level by a classical (separate) display.

What seems important to underline is that the two applications converge towards the same kind of structure of information, even if the two tools have been developed independently one of each other, and, even if the tools that complete the expert system are very dissimilar.

4.3. Philosophy of development

The two firms used contrasting philosophies of development :

- in the first one, the expert systems were developed with a strictly <u>technically-driven perspective</u>. The application running in the factory where the observations were gathered is as similar as possible to every application running in other factories of the group.
- in the second one, the expert systems (for each factory) were developed with the option of <u>maximising the consistency</u> between the human operator and the expert system assistant. Each application is thus different from all of the other applications developed in the industrial group.

In the "technically-driven" perspective :

- the only knowledge elicitation carried out in the factory was aimed at identifying the technical particularities of the process, in as much as a general description of the "best" control strategy has already been developed by engineers of the technical

service, starting from an engineer's point of view of the "best" way to control the process ;
- the global reasoning structure adopted by the expert system issues from this logical decomposition of the task carried out by engineers of the office (who don't really care about the details of the operators' activities) ;
- the terminology and inferential network included by the final tools fit the engineers' rationality (the same for every factory) instead of that of the (local) operators.

In the "maximal consistency" philosophy, the point of view was the opposite :

- an extensive knowledge elicitation was carried out in each factory with the operators (as was also the case for the factory where the observation were gathered) ;
- the global structure of reasoning included in the expert system is constructed based on the strategies used by operators (more precisely, by at least one operator) when manually controlling the process ;
- the terminology and inferential network used fit the local customs, and were not supposed to be transferable to another site.

- The intention put forward in this second approach was that the tool must always respect the "local" practices of the operators : their terms were used, their intermediary inferences and their own presentation of information were respected in the final tool. In accordance with the principle of classical Human Centred Automation presented in the introduction, this philosophy was explicitly chosen to optimise the acceptation of the tool and its appropriation by the operators and the team.

4.4. Expert system utilisation

In the case of the tools developed in the technically-driven perspective, the tools were nearly always used in closed-loop mode (about 90 % of the time) and in a similar manner by the operators.

For the other expert system, the disparity between operators concerning their rate of use of the expert system in closed-loop was very large. Some of them tried to use it all the time, others preferred to turn it to open-loop mode at the first sign of a discrepancy between what they would have done and what the expert system proposed to do. Moreover, after having turned the system to open-loop mode, they did not even use its diagnosis to support their own decision. Some of them even stopped considering the suggestion made by the tool.

So the utilisation of the expert system appears to be more significant in the case of the application developed in a technically-driven perspective than for the tool developed with the perspective of maximal consistency.

4.5. Trust in the reasoning capacities of the expert system

In the two factories, in certain circumstances, some operators present an excessive degree of confidence in the reasoning capacities of the tools. In both cases, operators could attribute a flexibility in the inference processes of the expert system that was not justifiable. When questioned about what could have occurred if, after a poorly adapted decision of the expert system, they had not turned it off, operators answered that "the system would have (at last) understood its mistake and corrected it" or that "it would not have driven the process in that bad direction for a long time". This kind of interpretations is, of course, completely erroneous, and illustrates some gaps in the mental model of the capacities of the expert system.

If in both situations some operators manifest this phenomenon of over-trust, it should be underlined that globally, the level of trust in the expert system is very different in the two situations. Many operators find if difficult to predict the behaviour of the expert system in the case of the tools developed in the "maximum consistency" orientation. Surprisingly, they express a strong disagreement with the way the tool handles the task. They said that they did not understand the principles of the intermediary inferences, that the reasoning did not seem completely rational to them.

In as far as the other system is concerned, such problems do not appear, and it seems that everybody accepted the way the expertise is handle by the system, even if they conducted the process in a rather different way.

So it seems that the acceptance and trust given to the tool is higher for tools developed in a technically-driven perspective than for one developed precisely with the objective of being accepted through a greater correspondence to the natural mode of reasoning of the operators. The reason for this is not to be found (or at least not only) in a difference in the qualities of the two tools, but seems to lie in the process of knowledge elicitation in itself (see discussion).

4.6. Organisational problems linked to the maintenance and up-dating of the tool.

The consequences of the introduction of an expert system goes beyond the strictly technical, cognitive and acceptation problems. It concerns also organisational problems, namely linked to the evolution of the application, at least when the supervised process evolves, but also when some weakness in the expert system are found.

In the industrial group where the maximal consistency option was chosen, there are as many different expert systems developed as they are different factories. So it is not surprising to see that the problem of the up-dating of the tool is managed at the local level of each factory. In the unit where the observations were gathered, these evolution were the responsibility of a small team composed of a software engineer and one or two of the operators. The first individual was in charge of the technical aspects of the development, and the others provided their expertise for the feeding of the knowledge base. Some difficulties arose from the incremental nature of these kinds of adaptations. As operators are working on shift change-over, to test and include a new rule in the knowledge base could take a rather long time, and it is not rare that all the operators are not informed of the modification. In the same way, it could also be the case that the documentation that explains the inference done by the expert system was not completely up-to-date.

A second problem (not observed) with this option is that when managers rotate between sites, it could be hard for them to understand the logic of their predecessor.

In so far as the technically-driven perspective is concerned, the maintenance and up-dating of the tools follow a completely different pattern. As the systems are built on a single model for all the factories of the group, every decision of modification must revert to the technical team responsible for the group. So, when a problem is detected in a site, a note must be faxed to the technical team of the group and the modifications, when accepted are transmitted to the other factories. The procedure is much longer, but homogeneity is guaranteed. When managers change, as the system are nearly identical everywhere, the adaptation problems are limited.

4.7. Training aspects.

A last important consequence of the philosophies of development is related to the training of the operators. As the distance between their spontaneous strategies and the strategy adopted by the expert systems is perceived as important, it is not astonishing to see that the group where a technically-driven perspective is adopted put a great emphasis in the training of their operators. Effectively, they must learn much more to build their own mental model of the expert system functioning. This training also contributes to the homogeneity of the knowledge, and is seen by the manager as a way to improve manual strategies as well.

The importance of these training aspects is not put forward in the other industrial group.

DISCUSSION

The efficiency of the two perspectives is rather contrasted. According to a larger number of criteria, the performances of the "technically-driven" approach were better than those of the "maximal consistency " approach :

- the acceptation of the tool by the operators, and their trust in the reasoning carried out by the systems was better where the technically-driven perspective has been adopted ;
- their understanding of the intermediary inferences of the tools was also better.

Furthermore, the advantages of this technical perspective goes beyond the human-system interaction, for instance to the organisational aspects : in case of turn-over in the management team, the adaptation of a manager to a new factory is much easier as the differences between the two systems are small.

It seems that one of the principal reasons for the poor level of performance of the system developed with a maximal consistency approach lies, unexpectedly, in the importance of the debate about the individual supervisory strategies of the operators. Adopting an expert system that fits the operators' strategies leads to the necessity of making a choice amongst these strategies, and (as far as the inter-individual differences were important) amongst the operators. So instead of favouring the adoption of the tools by all the operators, this philosophy of development provoked a negative reaction towards the tools for most of them.

On the opposite, the system developed following a technically-driven perspective starts with potential defaults that lead many ergonomists to prefer the other perspective. Namely, the fact that the process control modes are a priori very distant from those of all the operators could lead to a global rejection reaction. It seems that the quality of the design as well as the discernment of actions that facilitate the acceptation of the tool (namely the training policy) have largely compensated these handicaps. The better success of this application has to be related to the quality of the measure accompanying the implementation and the evolution of the tool. Moreover, one of the trump cards of this approach (in comparison with the other one) is that the expertise included in the tool, because it does not make any reference to that of the users, looks perfectly neutral. So the potential conflicts of influence and of knowledge related to the expert system are limited.

Another point is that the difficulties linked to the follow-up of an application developed with regard to the local habits, at least in industrial groups, where the managers rotate between the various sites, are really considerable. This fact underlines the advantages of systems developed in a technical perspective, which is shared amongst the sites. It also shows that a global approach is useful before determining design choices that have long-term consequences.

CONCLUSION

The comparison of the two systems illustrates that the advantages associated with the approach where a maximal consistency between the process control modes of the expert system and those of the operators are not real in every circumstance. It seems that when human expertise is rather multiform, when there is a risk of conflicts related to that expertise, adopting a (more) technically-driven perspective could be more efficient. The observations reported here underline the potential advantages of the technically driven perspective, as far as the integration of the tools in the organisational structure of the firm is carefully taken into account, and as far as accompanying measures are implemented.

It has also been shown that the picture in this domain is much more contrasted than it would at first appear. A default on a cognitive perspective could become an advantage where organisational aspects are concerned, and vice versa.

This result underlines also the necessity to carefully consider the advantages and inconveniences of the two design choices presented here before any real time expert system implantation. Furthermore it shows the necessity of a deeper theoretical reflection related to the nature and characteristics of the desirable matching between the characteristics of *any* tools and those of its users. A reflection in this direction has already been initiated by Vicente (1990).

REFERENCES

De Keyser, V. and Woods, D.D. (1989). Fixation errors in dynamic and complex systems. In : *Advanced systems reliability modelling (A.G.* Colombo and R. Micenta, Eds) , Kluwer Academic Press, Dordrecht, NL.

Grosjean V. (1994a). Ergonomie des aides à la conduite des processus industriels complexes : 1. Bilan bibliographique. Notes Scientifiques et Techniques, n° 117, I.N.R.S.

Grosjean V. (1994b). Ergonomie des aides à la conduite des processus industriels complexes : 2. Résultats d'une enquête concepteurs - utilisateurs. Notes Scientifiques et Techniques, n° 118, I.N.R.S.

Lehner, P.E. (1987). Cognitive factors in user/expert-system interaction. *Hum.Fact.* 29(1), 97-109.

Minault, S. and Penalva, J.M. (1993). Vers une approche coopérative de l'aide à l'opérateur d'exploitation. Journée d'Etude *"Explication et Coopération Homme-Machine"*. Bagnols, France.

Mitchell, C.M. and Miller, R.A. (1983). Design strategies for computer based information display in real-time control system. *Hum. Fact.*, 25, 353-369.

Moray, N. (1987). Intelligent aids, mental models, and the theory of machines. *Int. J. Man-Machine Studies.*, 27, 619-629.

Muir, B.M. (1987). Trust between human and machines, and the design of decision aids. *Inter. J. of Man-Machine Studies*, 27, 5/6, 527-540.

Norman, D.A. & Draper S.W. (Eds.) (1986) *User Centered System Design.* Lawrence Elbraun Associate. London.

Vicente, K.J. (1990). Coherence and correspondance-driven works domains : implication for systems design. *Beh. and Inform. Technol.* 9(6), 493-502.

Woods, D.D. (1986). Paradigm for intelligent decision support. In : *Intelligent Decision Support in Process Environment.* (E. Hollnagel et al., Eds.) Nato Series, Vol F21. Springer Verlag. Berlin.

MODELING CAR DRIVING AND ROAD TRAFFIC

P.H. WEWERINKE
Department of Applied Mathematics
University of Twente
P.O. Box 217, 7500 AE Enschede
The Netherlands

Abstract This paper describes a modeling approach to road traffic based on two model structures. The first model deals with individual driver behavior for the primary driving tasks. The second model describes the traffic process of a large area at a more global level as traffic flows. The result is a model that describes the detailed effects of task variables of interest on traffic performance and safety.
In addition, the level of experience and the adaptive human capability to traffic system developments is considered by comparing two modeling approaches: a system theoretic approach (based on an Extended Kalman filter) and a neural network approach.

1 Introduction

Many detailed factors determine road traffic capacity and safety. An important category can be indicated with human factors (involved in the majority of accidents); others are related to: vehicle dynamics, regulations, environmental conditions, advanced technology, etc. The challenging problem is to establish the effect of these factors on raod traffic performance (road capacity, traveling time and safety (probability of incidents and accidents). The difficulty here is to relate detailed factors to overall measures of performance and safety of the total traffic process.

In this paper the approach to solve this problem is based on modeling the road traffic process. At one hand this requires enough model detail to account for the effect of the many factors which influence traffic behavior. At the other hand, the model has to deal with the ultimate issues of interest: traffic performance and safety.
This implies a modeling approach starting with the detailed description of human behavior in the most important driver tasks: lane keeping, car following and overtaking. The result of these sub-models are measures of traffic speed as a function of traffic density and of safety. These measures can be used as input for a more global description of road traffic in terms of traffic flows. Finally these traffic flow models have to be aggregated to cover a substantial road traffic area.

A specific issue which is discussed into more detail is the driver's experience level. This is an important factor with respect to driver behavior and it's impact on traffic performance and safety. This human adaptive capability however, is also important in a developing traffic system.
Therefore, human learning is considered and modeled based on two approaches. The first one is a system theoretic approach. In that case learning is modeled as an adaptive estimation process of unknown model parameters. The second approach utilizes a neural network to describe adaptively the input-output behavior of learning the driving task.

2 Model analysis of driving tasks

The overall goal of car driving is to go from A to B in a certain way (safely, in a given time, etc.). The principal tasks derived from this are lane keeping, car following and overtaking slower vehicles, avoiding a collision with oncoming cars, based on visual cues of the outside world.

2.1 Lane keeping

Lane keeping is based on two primary visual cues: the inclination of (or distance to) the road side and the direction of an aimpoint. The first visual cue provides information about the lateral position y. The direction of an aimpoint ψ_a at distance d ahead is given by

$$\psi_a = \psi + y/d \tag{1}$$

Equation (1) shows that for large d, $\psi_a \approx \psi$ and for small d $\psi_a \approx y/d$; in other words: depending on the 'looking' distance ahead d, the driving task resembles more a relatively easy heading control task or a relatively difficult position control task. This can be illustrated by a simple root locus analysis (Figure 1).
Assuming that the driver is generating a steering wheel deflection δ proportional to the system output (i.e., ψ, y or ψ_a), the closed loop system dynamics can be visualized by the poles of the root loci. The root loci show that aimpoint control resembles more heading control or position control depending on the additional lead (zero at $-u/d$, with u the forward driving speed) corresponding to the implicit heading feedback. The value of d determines the trade-off between system stability and the bandwidth of the path mode. So d can be optimized and can be related to the driver's experience level.

Two approaches can be followed to model the lane keeping task: a conventional servo-system approach (see McRuer and Krendel, 1974) and the optimal control model approach (Kleinman et al., 1970, Wewerinke, 1989). The former is useful for simple tasks, the latter is more appropriate for more complex (multivariable) tasks involving perception, information processing, attention sharing, etc. For the lane keeping task the servo-system modeling approach is selected because of it's simplicity.

The hypothesis is that the driver behaves in such a way that a good closed loop system is obtained. The assumption can be made that the driver responds to the heading error, with a first order lag with time constant T_h (a typical value is 0.5 s). K_h is adjusted optimally.
An illustrative example is obtained from Wewerinke, 1994. For a driving speed of 30 m, $T_h = 0.5s$ and K_h optimal, a lane changing task is considered. The response to four step commands is shown in Figure 2 for a looking distance $d = 100\ m$. The stable response is lagging (90% of the command is realized after 5 seconds) because of the dynamics of the vehicle, of the driver and primarily because of the control strategy (looking distance ahead d). This result is compared with a neural network approach in the following.

2.2 Car following

Car following is based on the visual cues α and β shown in Figure 3. A geometric analysis reveals that these cues are inversely proportional to the distance to the preceding car (X). Linearization of these relationships around a nominal condition yields a linear relationship between the visual cues and X (and between their derivatives) and u (the relative speed between the cars).

Also for the car following task a simple servo-system model is considered. Again a driver model is assumed consisting of a first order lag with a time constant $T_h = 0.5$ seconds.

The closed loop transfer function is given by

$$\frac{X(s)}{X_d(s)} = \frac{K_h K_x}{s(T_u s+1)(T_h s-1) + K_h s + K_h K_x} \quad (2)$$

The resulting closed loop poles consist of a well-damped second order speed mode and a first order path mode with a bandwidth of approximately K_x.

2.3 Overtaking

2.3.1 Task analysis

Overtaking is the most complex driving subtask including observing, information processing, decision making, planning, maneuvering and other traffic. The overtaking situation is shown in Fig. 4 for a two-lane road. The analysis for freeways is similar.

Assume that car i intends to pass the preceding car j at distance X_j ahead. The possibility to do so depends on the oncoming cars k and ℓ. At time $t_0 = 0$ car i decides at distance X_j from car j to accelerate from speed u_j to speed u_m (with time constant T_u. The distances to cars k and ℓ (with speed u_ℓ) are X_k and X_ℓ, respectively. At time t_1 car i changes to the left lane. At time t_2 car j is overtaken and car i returns to the right lane again.

The two conditions which must be satisfied are that at t_1 car k must be passed and car ℓ must be away far enough to allow the overtaking maneuver. This implies the two inequalities

$$X_k < X_j + (u_\ell + u_j)t_1 - S_k \quad (3)$$

$$X_\ell > X_j + S_j + S_\ell + (u_\ell + u_j)t_2 \quad (4)$$

with S_j the distance which car i is overtaking in the left lane with respect to car j and S_k and S_ℓ the safety distances to car k and ℓ, respectively.

Useful approximations can be derived (Hilberink, 1994) for t_1 and t_2 (for t_1 and t_2 smaller than T_u). The result is

$$t_1 \approx \sqrt{\frac{2.7 X_j T_u}{u_m - u_j}} \quad \text{and} \quad t_2 \approx \sqrt{\frac{2.7(X_j + S_j)T_u}{u_m - u_j}} \quad (5)$$

Combining eq. (5) with (3) and (4) shows that the decision criteria (3) and (4) depend on X_j (which can be identified with the overtaking strategy), car dynamics (T_u), speeds and safety margins.
The required gap between car k and car ℓ $X_\ell - X_k$)

can be determined as a function of X_j. The result in Fig. 5 shows that the required gap can be substantially reduced for a small X_j. So the possibility to trade-off $X_\ell - X_k$ and X_j allows car i to optimize its overtaking strategy depending on the momentaneous traffic situation.

2.3.2 Overtaking model

The speed of car i is represented by a first order process driven by a commanded speed (u_c) and a first order disturbance input (w_u). Other traffic is modeled with a constant average speed. The specific speed of each car is unknown to car i and has to be estimated based on the observed (derivative of) distances to the other cars.
The system model of the total process can be expressed in the general form

$$X(k+1) = f(X(k), U(k), W(k)) \quad (6)$$

with $X = col(W_{u_i}, u_i, X_j, X_k, X_\ell, X_{ik}, X_{i\ell})$ with X_{ik} and $X_{i\ell}$ the adjoint states representing the inequalities (3) and (4), combined with (5).

The distances to other cars can be estimated based on the visual angles shown in Fig. 3. The nonlinear relationships can be expressed in the general form

$$y_p(k) = g(X(k-i)) + v(k-1) \quad (7)$$

with v the observation noise and i the time delay. By means of an Extended Kalman filter the states can be estimated. The estimated variable X_{ik} and $X_{i\ell}$ are used to decide whether (D_1) or not (D_0) overtaking is possible, in case the distance to the slower preceding car becomes too small. This involves a sequential decision process to determine continuously the mode of car i: 1. driving with speed u_c, 2. car following with speed u_j and 3. overtaking, which is basically a pre-program maneuver (accelerating at t_0 till u_m and lane changes at t_1 and t_2).

In Table 1 the overtaking performance is summarized as a function of traffic density, for a given car ($T_u = 10\ s$) and overtaking strategy ($X_j = 10\ m$). Other traffic has a speed of 20 m/s (45 m/h). The desired speed of car i is 30 m/s (67.5 m/h). The density in the right lane (d_r) and in the left lane (d_ℓ) are varied independently. A density of 0.005 (average number of cars per meter with a Poisson distribution) results in a relative velocity (v_r) increase of 5.2 m/s (11.7 m/h) and 181 overtakings (in a simulation period of 2 hours). Especially an increase in the left lane traffic density reduces the performance quickly; for $d_\ell = 0.02$ hardly any overtakings are still possible.

2.4 Traffic flow modeling

The description of large traffic areas requires a more global modeling approach. In that case traffic behavior can be described in terms of traffic flows, instead of individual cars. This involves the description of average measures, such as speed, density and through put, per road segment.
The situation is shown in Fig. 6. For the traffic density ρ can be written

$$\rho_i(k+1) = \rho_i(k) + c[q_{i-1}(k) - q_i(k) + r_{i-1}(k) - s_i(k)] \quad (8)$$

with q the through put (number of cars per unit time), r and s the traffic that enters and leaves the road, respectively, and c a constant. Furthermore, it will be clear that

$$q = \rho v \quad (9)$$

with v the average speed per road segment. So combining (9) with (8) yields a recursive equation for ρ, which can be solved if $v(\rho)$ is known.

Traditionally this relationship between v and ρ is established empirically based on available data: typically, the speed decreases rapidly around 50 cars/km and is close to zero beyond 100 cars/km. However, such an empirical basis excludes the possibility to predict the effect of detailed factors on the traffic process.
To solve this problem the traffic flow model is combined with the foregoing detailed driver models. One of the outputs of these models is the average speed as a function of traffic density (see before) and this is precisely the required input for the traffic flow model.

The result is a model structure that describes the relationship between detailed traffic factors (related to driver behavior, traffic system design, environment, etc.) and overall measures of road capacity and safety.

3 Adaptive driver behavior

An important aspect of car driving is the level of experience and the adaptive human capability to traffic system developments.
Learning the driving task can be related to the following task aspects.
Knowledge has to be build up of the car response to control inputs (steering angle, gas, brakes). This internal model of the vehicle dynamics plays an important role in decision making, planning and controlling.
The second task aspect that depends on the level of experience is the use of the visual cues. This concerns the specific cues involved, the accuracy with which the cues are perceived and the attention allocation among the various cues.

Thirdly, the level of experience can be related to the control strategy of the lane keeping task and the car following task. This implies trade-offs between the multiloop feedbacks (heading-, speed- and position feedbacks) but also a possible feedforward control based on the forward (curved) road. The perceived course of the forward road segment can be accounted for by means of a preprogrammed control sequence to track the next road segment. However, for most road situations, it seems reasonable to assume exclusively feedback control.
The decision to overtake is based on the relationships between the distances to the oncoming cars and the preceding car. The experience level of the driver can be related to an optimal passing strategy based on these relationships and on the time coordination involved.

Two approaches are followed to model the aforementioned learning-related issues. The first one is based on system theory. Learning is modeled as an adaptive estimation process of unknown model parameters based on new data (experience), related to the vehicle dynamics, the visual cues, the lane keeping control strategy and the overtaking maneuver.
An important question in this context is what one's initial knowledge is of the system. One possible assumption is that a beginning driver knows nothing.
The model can be used to predict the effect of learning for the track keeping task, the car following task and the overtaking task in the case of other traffic. The task configurations considered include the variation in average traffic speed, visibility conditions and traffic density. Model predictions are in terms of performance measures and measures which are related to traffic safety, such as probabilities of collisions and decision and safety margins.

In addition to this modeling approach, learning the car driving task is described in terms of a neural network (NN). Basically the NN model is an input-output model of the human operator, relating the task inputs (visual cues) to the control outputs (steer, gas and brakes). The assumed model structure is a feedforward network with three layers (one hidden layer). The model parameters are the connection weights and thresholds, representing the memory of the network.
A backpropagation scheme is assumed to adjust the weights to mimic human learning. For that purpose input-output data of the afore-mentioned model are used. The backpropagation is aimed at the minimization of the output error (defined as the difference between the simulated outputs and the corresponding NN outputs).

In order to compare both approaches describing human learning behavior in car driving, the preliminary simulation results of the lane keeping task are presented in the next section.

4 Simulation results

As a first step the lane changing maneuver is considered, of which the system model results are discussed in 2.1 and which is shown in Fig. 2.
The NN was trained by presenting the system model results of 1000 lane changes between (randomly) 0 and 2 meter at intervals between 0 and 10 seconds. The heading and lateral deviation were presented as inputs; the steering wheel deflection was the output of the NN. Next, the same commanded lane changes as shown in Fig. 2 (every ten seconds a lane change of 1 meter) were presented to the NN. The resulted response shows a rather close agreement with the system model results.

Next, the heading control task was considered, in the presence of a random disturbance. The NN was trained on the basis of 25 seconds of data (250 data pairs). Fig. 7 shows the results for a varying number of iterations to adjust the weights of the NN. After about 100 iterations a reasonable close agreement with the system model results is obtained, showing that the NN is able to learn the relationship between heading and steering wheel deflection (the driver model) in spite of the presence of the random disturbance.
After teaching the NN the heading task based on 25 seconds data it is investigated how well the NN has learned the task (for new data). This is shown in Fig. 8 for for 200 seconds. The NN (dashed line) matches the model heading closely (apart from the extreme values).

The second approach to describe learning is formulated as an adaptive estimation process of unknown parameters.In the present example 2 parameters (k and a) are related to the heading dynamics and two parameters (K_h and T_h) are part of the driver model. Two cases are considered. For case 1 it is assumed that K_h and T_h (representing the driver) are unknown. The Extended Kalman filter results are shown in Fig. 9. After about 50 seconds the heading response is learned. For the second case all four parameters are assumed to be unknown. This is comparable to the NN which has also no information about the process structure. So learning has to be based solely on the input-output data. The EKF results are shown in Fig. 10. Now it takes a rather long time to learn the heading response (say 400-500 seconds). So, one could be tempted to conclude that the NN is in this ease superior. However, one must keep in mind that the NN needed many iterations based on the data of 25 seconds.

5 Concluding remarks

In this paper a modeling approach to describe road traffic is based on two model structures. The first model deals with individual driving behavior for the primary driving tasks so as to describe the detailed effect of task variables of interest on traffic performance and safety. The resulting average speed can be assessed as a function of traffic density.
This is the input for a second model which describes the traffic process of a large area at a more global level as traffic flows. The result is a model structure that can be used to predict the effect of detailed traffic factors on overall measures of road capacity and safety.

In addition, the level of experience and the adaptive human capability to traffic system developments is considered by comparing two modeling approaches: a system theoretic approach (based on an Extended Kalman filter) and a neural network (NN) approach.
Preliminary simulation results show that a NN is able to learn the driving task considered. Also an Extended Kalman filter was able to learn the task, requiring more or less data depending on the a priori knowledge of the process. These two issues determine which of the two approaches is the best to describe human learning behavior. This has to be checked out in an experiment with real human operators. This is planned as the next step.

References

[1] D.T. McRuer and E.S. Krendel, *Mathematical models of human pilot behavior*, AGARDograph, Nr. 188, 1974.

[2] D.L. Kleinman et al., *An optimal control model of human response, Part I: theory and validation*, Automatica, Vol. 6, 1970.

[3] P.H. Wewerinke, *Models of the human observer and controller of a dynamic system*, Ph.D. thesis, University of Twente, 1989.

[4] P.H. Wewerinke, *Modeling human learning involved in car driving*, Proc. IEEE SMC, San Antonio, 1994.

[5] H. Hilberink, *A comparison of system theoretic models and neural networks applied to human learning involved in car driving*, Thesis (in Dutch), University of Twente, 1994.

d_l \ d_r	0.05	0.1	0.2
0.05	5.2 (181)	3.8 (222)	2.9 (282)
0.1	3.0 (101)	1.7 (98)	1.1 (105)
0.2	0.3 (12)	0.2 (14)	0.1 (11)

(·) number of overtakings

Table 1: Overtaking performance (v_r) as a function of traffic density

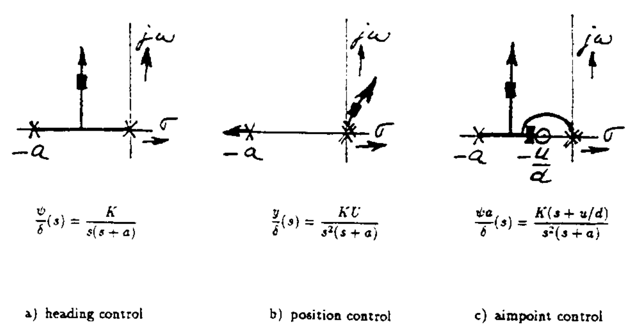


Fig. 1. Root loci of the lateral control modes

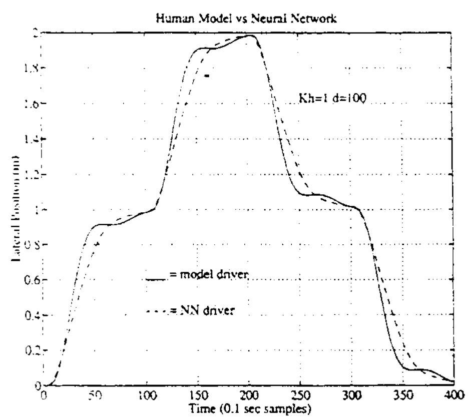


Fig. 2. Lane keeping performance for the two models

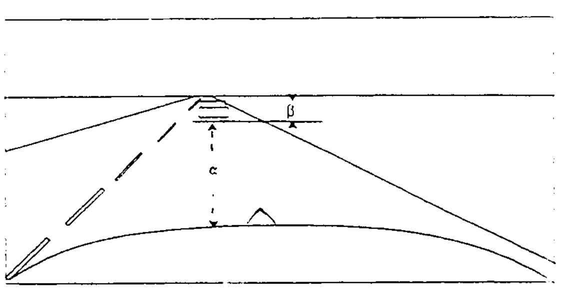


Fig. 3. Visual cues car following

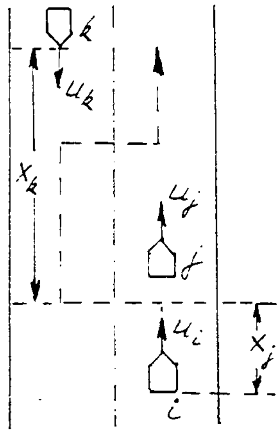


Fig. 4. Overtaking task

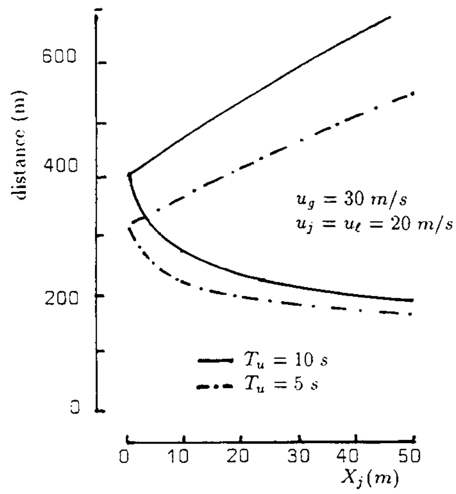


Fig. 5. Relationships between distances to other cars

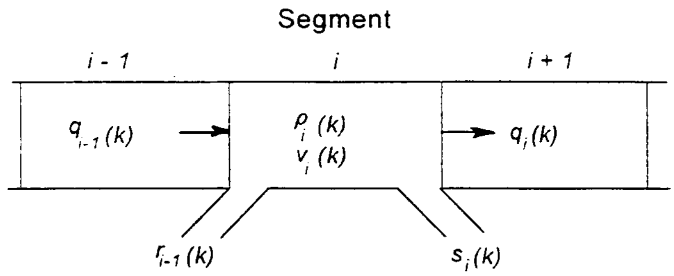


Fig. 6. Traffic flow modeling

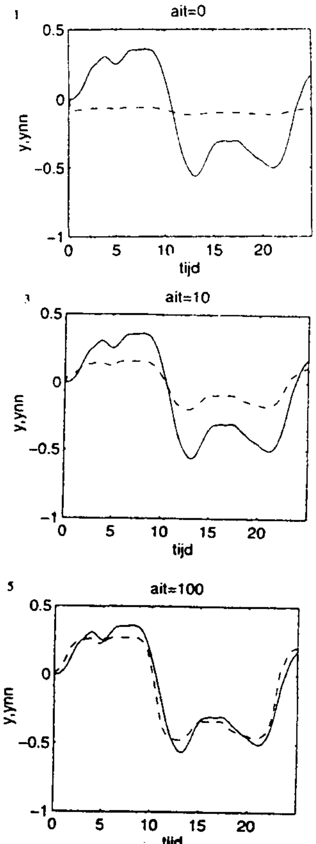


Fig. 7. The NN learning the heading control task

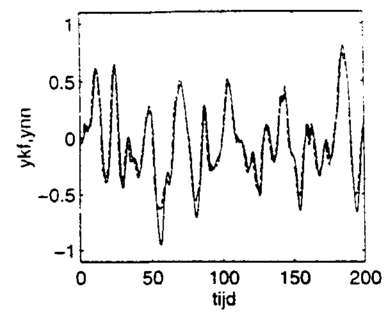


Fig. 8. The agreement between the NN and the driver model for new data

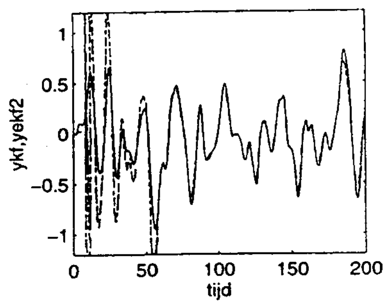


Fig. 9. The EKF learns the heading task for two unknown parameters

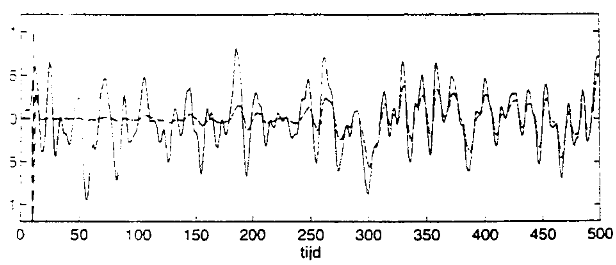


Fig. 10 The EKF learns the heading task for 4 unknown parameters

AN ESTIMATION OF THE HAZARD-CONTROLLABILITY OF DRIVER-SUPPORT SYSTEMS

Yoshinobu Sato*, Eiichi Kato and Koichi Machida****

* *Dept. of Electronic and Mechanical Engg., Tokyo Univ. of Mercantile Marine 2-1-6 Etchujima, Koto-Ku, Tokyo 135*
** *Technical Research Laboratory, Hino Motors, Ltd., 3-1-1 Hinodai, Hino-Shi, Tokyo 191*

Abstract: Advanced safety vehicles are supposed to materialize driver-support systems such as distance-warning systems and emergency brake systems. In the present paper, the hazard-controllability of the driver support systems is estimated based on deductive methodology. First of all, hazards are identified and the targeted systems are defined. Next, accidents causation models are developed in terms of fault-trees with the hazards being top-events. Then the accident causation models are analyzed qualitatively and quantitatively. Finally the effectiveness of the driver support-systems is estimated by comparing the statistics of the hazards for both vehicles equipped and those not equipped with the driver-support systems.

Keywords: Vehicle, Automobile, Man/machine systems, Human error, Alarm systems, Back-up systems, Fault-tolerant systems, Risk, Safety analysis

1. INTRODUCTION

A distance-warning system (**DWS**) measures the distance between one's vehicle and the vehicle in front by means of laser beams and gives warning to a driver if his vehicle approaches the preceding one excessively. **DWS** has been developed and has lately been put to practical use by trucks and busses to prevent rear-end collision in Japan. An emergency-brake system (**EBS**) stops a vehicle automatically before it collides with the preceding vehicle. **EBS** is now in contemplation.

It is not only interesting but also important to analyze the hazard-controllability of driver support systems (**DSSs**) such as **DWS** or **EBS** for optimization of system structures and system safety assessment.

Two opposite approaches exist for hazard-controllability analysis; one is inductive and the other is deductive. The former can be applied to systems with long operation histories and thus enough experiences of system abnormalities. This approach statistically analyzes large numbers of accidents, incidents and other occurrences observed so far. On the other hand, the deductive approach deals with relatively innovative systems where historical data are few. This approach first establishes causal models where less experienced events are successively expressed as functions of more experienced events. Ultimately they are expressed as function of basic or axiomatic events having enough statistical data. It has been, for example, applied to the studies on robot safety (Kandel, *et al.*, 1988; Sato, *et al.*, 1987; Sato, *et al.*, 1990).

The deductive approach consists of the following fundamental stages:
1. Hazard identification and planning of hazard-restraint structures,
2. Development of causal models,
3. Analyses of causal models, and
4. Estimation of the hazard-controllability of the hazard-restraint measures.

The data base for analyzing effectiveness of **DSSs** statistically has not been available. We, therefore, try to analyze and estimate the hazard-controllability of **DSSs** based on the deductive approach.

Here, notation and terminology follow those given by Henley (1992).

Abbreviations

BHEP: basic human error probability
DSS: driver-support system
DWS: distance-warning system
EBS: emergency-brake system
ENO: statistically expected number of occurrences of a top event
KITT: kinetic tree theory
THERP: technique for human-error rate prediction

2. DRIVER SUPPORT SYSTEMS (DSSs)

2.1 Distance-warning system (DWS)

The **DWS** consists of such main components as a man-machine interface, processor, laser sensor, speed sensor, self-diagnosis monitor as shown in Figure 1. The functions of these components are:

1. The laser sensor measures the distance from the vehicle with DWS to the preceding one and the speed sensor senses its own vehicle's speed, respectively.
2. The processor calculates safety distance from the vehicle to the preceding one and compares actual distance with the safety distance taking account of the inputs from the laser sensor and the speed sensor. It gives outputs to the man-machine interface if necessary.
3. The processor diagnoses the sensors continuously and gives the results to the man- machine interface and the self-diagnosis monitor.
4. The man-machine (**M-M**) interface has a digital distance display (which indicates the distance from the vehicle to the preceding one in real time), an audio and photo-display (which notices and warns the driver of an excessive approach), a lamp (which indicates the windowpanes of laser sensor being clouded up with dust), a **DWS**-failure indicator (which lamps if the processor diagnoses a failure of **DWS**), safety-distance setting mode buttons, a power switch, etc.
5. The self-diagnosis monitor indicates the failure mode on demand.

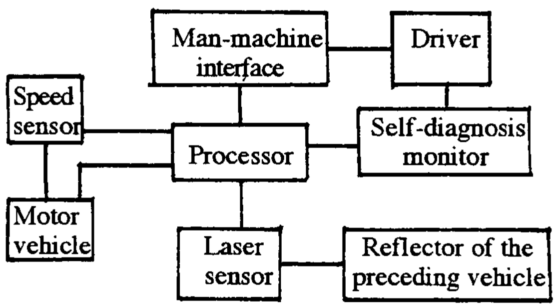


Fig. 1 DWS system block diagram

2.2 Emergency-brake system (EBS)

Figure 2 shows the **EBS** block diagram consisting of the same components as **DWS** except for the man-machine interface, power switch, actuator and brake mechanism. The actuator acts the brake mechanism in accordance with outputs from the processor to control the vehicle's speed. A driver can release the **EBS** using the power switch.

3. HAZARD IDENTIFICATION AND HAZARD-RESTRAINT STRUCTURES

We first identify the following hazards:
Primary hazard: A vehicle, which is not equipped with any **DSS**, collides with the vehicle in front because of a driver's error such as looking aside or dozing.

A vehicle can be equipped with a **DSS** to control the primary hazard. However, there could still be the possibility of the following hazards:
Secondary hazard: A vehicle equipped with a **DSS** collides the vehicle in front because of driver's error and/or a malfunction of the **DSS**.

4. DEVELOPMENT OF CAUSAL MODELS

We describe the causality of an accident, which is brought about by a hazard, with fault trees. Top events of fault trees represent system failure modes resulting in injury or damage; accordingly, system failure modes are determined by hazards. The primary and secondary hazards identified in the previous section result in the system failure mode "an outbreak of a rear-end collision".

4.1 Causal model for the primary hazard

The system failure mode "an outbreak of a rear-end collision" (**E**: the event identification code, and so forth) occurs when the event, a driver fails to decelerate (E0,0), arises under the condition that "the speed of the driver's vehicle is greater than the speed

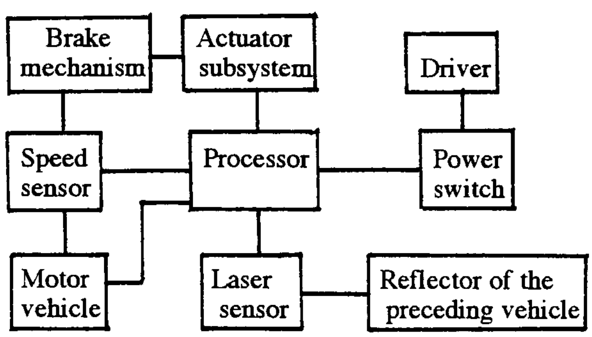


Fig. 2 EBS system block diagram

of the preceding one"(C). Event E0,0 occurs when one of the three following events, dozing (E2,1), wrong handling (E2,2), and wrong judgment (E2,3) arises. Here, it is assumed that the driver's vehicle is not equipped with a **DSS** (E1,1). The system failure logic for the primary hazard is represented by the fault tree in Figure 3.

4.2 Causal model for the secondary hazard created by a vehicle with DWS

The secondary hazard created by a vehicle with DWS could bring about the accident, an outbreak of a rear-end collision (**E'**). In this case, the top event **E'** also occurs when the event, a driver fails to decelerate (E0,0'), arises under the condition that "the speed of the driver's vehicle is greater than the speed of the preceding one" (**C'**) as shown in Figure 4.

We can divide event E0,0' into two classes, i.e., "under malfunction of **DWS**" (E1,2) and " under function of **DWS**" (E1,3) as shown in Figure 4. Here, the event "malfunction of DWS" means that the **DWS** does not generate proper notice and warning.

Event E0,0' occurs when one of the three events, dozing (E2,1'), wrong handling (E2,2'), and wrong judgment (E2,3') arises, under the condition that "the **DWS** is functioning" (E1,3). Event E0,0' also occurs when two events, malfunction of DWS (E2,4), driver's error (E2,5), exist simultaneously.

Event E2,4 is classified into two cases "under **DWS** power switch off"(E3,1) and "under DWS power switch on"(E3,2). Event E2,5 arises when one of the three events, dozing (E2,1), wrong handling (E2,2), and wrong judgment (E2,3) exists, under the condition that the **DWS** is malfunctioning. Event E3,1 arises when the event "a driver forgets toswitch on"(E4,1) or "a driver keeps switch off"(E4,2) exists. And event E4,2 is caused by the event, wrong warning (E5,1) or warning too frequent (E5,2).

Event E3,2 occurs when one of the five events, laser-sensor failure (E4,3), speed sensor failure (E4,4),processor failure (E4,5), man-machine interface failure (E4,6), and "power is not supplied to the processor from the motor vehicle"(E4,7) arises. Events, E4,5, E4,6 and E4,7, which are described using diamond event symbols, are further developed to more basic faults in order to obtain sufficient statistical data. However, we have to stop the description because of the paper space.

4.3 Causal model for the secondary hazard created by a vehicle with EBS

The secondary hazard created by a vehicle with **EBS** could bring about the accident, an outbreak of a rear-end collision (**E**"). The causation of this accident is similar to that discussion in the previous section as shown in Figure 5. However the following logic is different:
Event "failure to deceleration" (E0,0") arises when two events "malfunction of EBS" (E1,4) and "driver's error" (E1,5) exist, simultaneously. Event E1,4 is divided into two cases "under power switch off" (E2,6) and "under power switch on" (E2,7).

Event E2,6 occur when one of the events, "forget to switch on" (E4,8) and "keep switch off" (E4,9) because of "wrong functioning" (E5,3).

Event E2,7 arises when one of the events "laser sensor failure" (E4,10), "speed sensor failure" (E4,11), "processor failure" (E4,12), "actuator failure" (E4,13), "power not supplied" (E4,14) and "brake mechanism failure" (E4,15) occur.

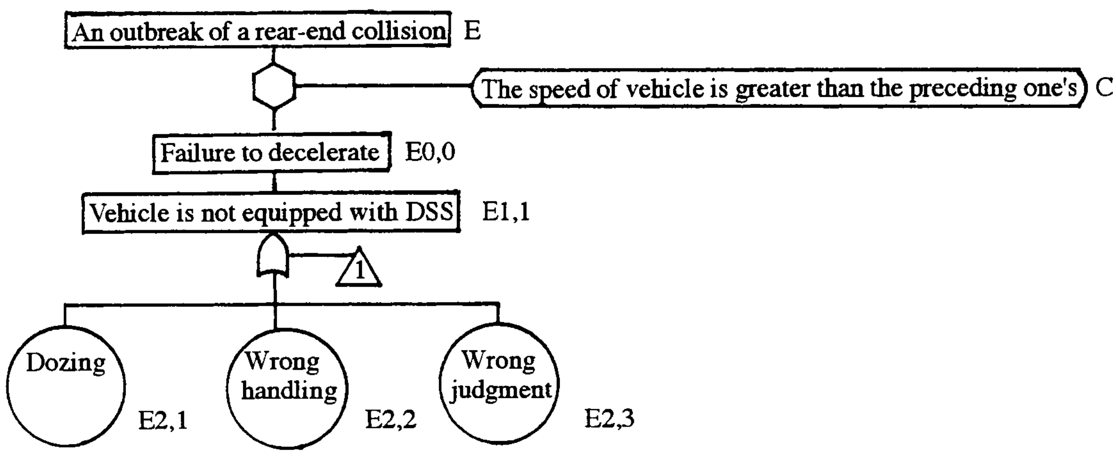


Fig. 3 A fault tree for the primary hazard

5. ANALYSIS OF CAUSAL MODELS

The statistically expected number of occurrences of the top-event (***ENO***) is the main dependability for risk assessment. The methodology for quantifying the failure logic and estimating the ***ENO*** has been developed by Vasely (1970), Sato (1986), etc. Here, we examine the ***ENO*** using short-cut calculation methods derived from **KITT** (Henley, et al., 1992).

KITT first evaluate the probability of each minimal cut set, i.e., the probability of simultaneous existence of basic events which compose the minimal cut set. The summation of the probabilities of minimal cut sets approximates the probability of event E0,0 (or E0,0' or E0,0") if each probability of the minimal cut set is sufficiently smaller than unity. Thus, we can easily calculate the ***ENO*** by integrating the unconditional failure intensity of the system over inspection interval ***T***.

5.1 Quantification of basic events

The following specification of (basic) events and assumptions are reasonable:

1. Event C (C', C"), the speed of vehicle is greater than the preceding one's, appears on the average of n times an hour.
2. The probabilities of events E2,1 (E2,1'), E2,2 (E2,2'), E2,3 (E2,3'), E4,1 (E4,8), and E4,15, per demand, are p_1 (p_1'), p_2 (p_2'), p_3 (p_3'), p_4 , and p_5 respectively.
3. The occurrences of events E5,1, E5,2, E4,3 (E4,10), E4,4 (E4,11), E4,5 (E4,12), E4,6, E4,7 (E4,14), E5,3 and E4,13 can be modeled by non-repairable exponential distributions with failure rates λ_1, λ_2, λ_3, λ_4, λ_5, λ_6, λ_7, λ_8, and λ_9 [hour^{-1}], respectively.
4. The inspection interval of the system is T [hours], and the basic events do not exist immediately after a inspection.

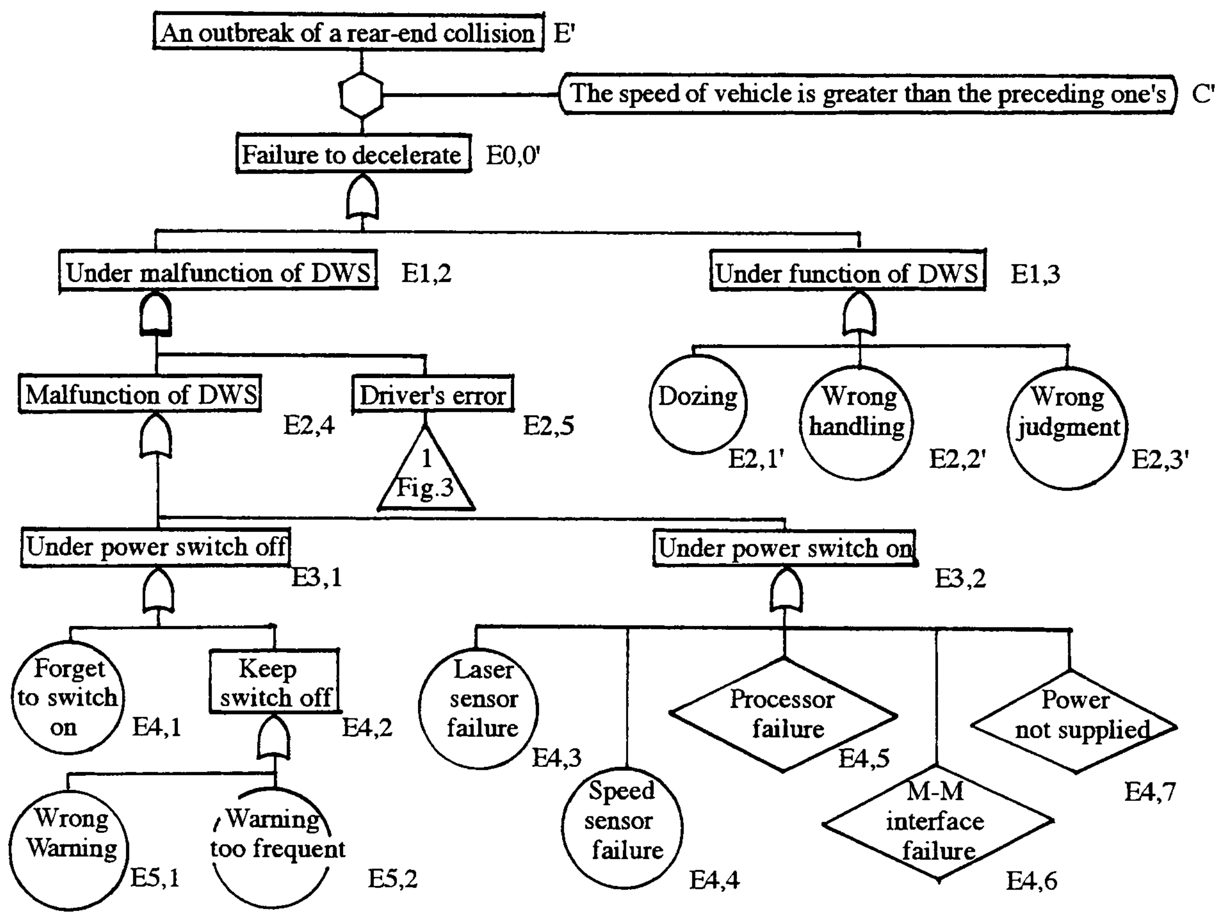


Fig. 4 A fault tree for the secondary hazard created by a vehicle with DWS

5.2 *Statistically expected number of occurrences of top event*

We define ENO_1 as the ENO for the primary hazard during one inspection interval, then from minimal cut sets and the quantified basic events, ENO_1 is:

$$ENO_1 = n\,(p_1 + p_2 + p_3)T \qquad (1)$$

Similarly, ENO_2 for the secondary hazard created by a vehicle with DWS is:

$$ENO_2 = n\,[p_1' + p_2' + p_3' + \{p_4 + (\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 + \lambda_7)T/2\}(p_1 + p_2 + p_3)]T \qquad (2)$$

Here, we have used the following relationship:

$$Pr\{\text{a hardware with failure rate } \lambda \text{ is at fault at time } \tau\} = 1 - \exp\{-\lambda\tau\}$$

$$\fallingdotseq \lambda\tau \quad (\ll 1)$$

In the same way, ENO_3 for the secondary hazard created by a vehicle with **EBS** is:

$$ENO_3 = n\,\{p_4 + p_5 + (\lambda_3 + \lambda_4 + \lambda_5 + \lambda_7 + \lambda_8 + \lambda_9)T/2\}(p_1 + p_2 + p_3)T \qquad (3)$$

6. ESTIMATION OF HAZARD-CONTROLLABILITY OF DSSs

It is reasonable to define the ratio of ENO_1 to ENO_2, or of ENO_1 to ENO_3 as the criteria for thehazard controllability of **DSS**. Then, we obtain the following criterion from equations (1) and (2):

$$\eta_1 = ENO_1 / ENO_2$$

$$= (p_1 + p_2 + p_3)/[p_1' + p_2' + p_3' + \{p_4 + (\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 + \lambda_7)T/2\}(p_1 + p_2 + p_3')] \qquad (4)$$

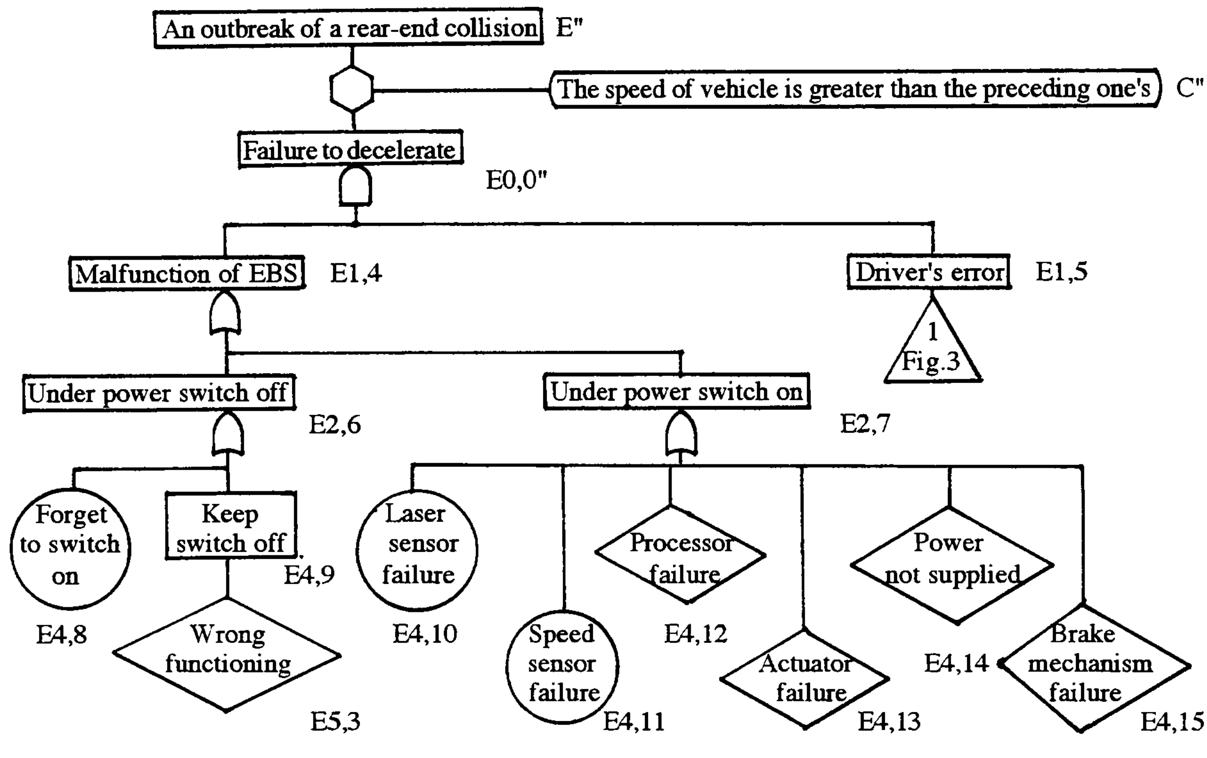


Fig. 5 A fault tree for the secondary hazard created by a vehicle with EBS

From equations (1) and (3),

$$\eta_2 = ENO_1 / ENO_3$$

$$= \{p_4 + p_5 + (\lambda_3 + \lambda_4 + \lambda_5 + \lambda_7 + \lambda_8 + \lambda_9)T/2\}^{-1} \quad (5)$$

We now evaluate criteria η_1 and η_2 by calculating equations (4) and (5). The specific parameters of the systems are as follows:

1. The basic human error probabilities (**BHEPs**) for the human errors, "dozing", "wrong handling" and "wrong judgment" are 10^{-5}, $5x10^{-4}$ and $6x10^{-3}$, respectively (Swain, *et al.*, 1980). We suppose that a driver has five chances to correct his error on the way to a collision, i.e., an event tree with five branches models the process of collision, based on **THERP**. If we put the transition probabilities of "dozing", "wrong handling" and "wrong judgment" as 0.2, 0.02, and 0.02 (which are based on subjectivistic viewpoints), then the following values are obtained: $p_1 = 1.6x10^{-8}$, $p_2 = 8x10^{-11}$ and $p_3 = 9.6x10^{-10}$
2. When a **DWS** is malfunctioning, the human error probabilities are equal to those in a system not equipped with any kind of **DWS**. When a **DWS** is functioning, we put the transition probabilities of "dozing", "wrong handling" and "wrong judgment" as 0.002, 0.01 and 0.01, respectively. Then, the probabilities of events, E2,1', E2,2' and E2,3' become $1.6x10^{-16}$, $5x10^{-12}$ and $6x10^{-11}$, respectively.
3. The maximum and minimum of the probability of event E4,1 or E4,8 are 0.1 and 0.01, respectively.
4. Hardware failures are: $\lambda_1=5x10^{-7}$, $\lambda_2=6.7x10^{-7}$, $\lambda_3=5x10^{-5}$, $\lambda_4=5x10^{-5}$, $\lambda_5=5x10^{-6}$, $\lambda_6=5x10^{-5}$, $\lambda_7=5x10^{-6}$, $\lambda_8=10^{-5}$ and $\lambda_9=5x10^{-6}$[hour^{-1}] and $p_5 = 10^{-6}$ [demand^{-1}]. These values are based on reliability data books (Henley, et al., 1992; IEEE, 1983).

Let inspection interval T be $2x10^3$ [hours], then the following results for η_1 are obtained:

$$\eta_{1max} = 5.8 \quad (\text{at } p_4 = 0.01) \quad (6)$$

$$\eta_{1min} = 3.8 \quad (\text{at } p_4 = 0.1) \quad (7)$$

For η_2

$$\eta_{2max} = 7.7 \quad (\text{at } p_4 = 0.01) \quad (8)$$

$$\eta_{2min} = 4.5 \quad (\text{at } p_4 = 0.1) \quad (9)$$

7. SUMMARY AND CONCLUSIONS

In the present paper, we estimate the hazard-controllability of the driver-support systems (**DSS**), i.e., a distance warning system (**DWS**) and an emergency-brake system (**EBS**) for motor vehicles based on deductive methodology. First, we identify hazards and define the systems. Next, we develop fault-trees with the top event, occurrence of a rear-end collision, for three kinds of vehicles: a vehicle not equipped with any **DSS**, a vehicle with a **DWS** and a vehicle with an **EBS**. Then we analyze the fault-trees using the **KITT** (kinetic tree theory) and **THERP** (technique for human-error rate prediction), and estimate the statistically expected numbers of occurrences of the top event (i.e., a rear-end collision) for these cases. Finally, the effectiveness of the **DSS** are estimated by introducing criteria which can make a comparison among these statistics. The estimation concludes that motor vehicles with **DSS** have a relative risk of rear-end collisions significantly lower than those without **DSS**. In order to take drastic steps, it is strongly suggested that a vehicle should be equipped with both **DSSs**, i.e., a **DWS** and a **EBS**.

REFERENCES

Henley, E. J. and H. Kumamoto (1992). *Probabilistic Risk Assessment*, IEEE Press, Inc., New York.

IEEE, Std. 500 (1983). *Reliability Data*, IEEE Press, Inc., New York.

Kandel, A., E. Avni (1988). *Engineering Risk and Hazard Assessment*, pp.61-80, CRC Press, Inc., Boca Raton, Florida.

Sato, Y., K. Inoue and H. Kumamoto (1986). The safety assessment of human-robot systems, *Bull. JSME*, **Vol.29**, pp.3945-3951.

Sato, Y., K. Inoue and H. Kumamoto (1987). The safety assessment of human-robot systems; *JSME Int. J.*, **Vol.30**, **No.260**, pp.350-356.

Sato, Y., E. J. Henley and K. Inoue (1990). An action-chain model for the design of hazard-control systems for robots, *IEEE Trans. on Reliability*, **Vol.39**, **No.2**, pp.151-157.

Swain, A. D., et al. (1980). *Handbook of human reliability analysis with emphasis on nuclear power plant application*, NUREG/CR-1278.

Vasely, W. E. (1970). A time-dependent methodology for fault tree evaluation, *Nuclear Engg. and Design*, **Vol.13**, p.337.

DAISY - A DRIVER ASSISTING SYSTEM WHICH ADAPTS TO THE DRIVER

J.P. Feraric', R. Onken

University of the Armed Forces
Institute of System Dynamics and Flight Mechanics
Werner-Heisenberg-Weg 39
D - 85577 Neubiberg, Germany

Abstract: DAISY is a model based Driver AssIsting SYstem which includes a component that accounts for the adaptation to the individual driver. Artificial neural networks based on the ART architectures were used to realize this module. Besides a brief survey of the main features of DAISY this paper emphasizes the modeling of the driving behaviour with ART networks. The way the ART networks are adapted to the problem is described. This includes the realization of a hierarchical ART structure and of a situation specific weighting of the components of the feature vector. The network parameters were adjusted through Genetic algorithm optimization.

Keywords: Driver models, Driver behaviour, Driver assistance, Adaptive sytems, Neural networks, Fuzzy ART, Genetic algorithms, Realtime learning system

1.DAISY ARCHITECTURE

1.1 Background

In the industrialized countries the individual mobility is an essential part of the modern culture. However individual flexibility becomes more and more limited right now through high traffic density resulting in bumper to bumper traffic situations and an increasing number of accidents. Traffic research has found that a very high percentage of accidents results from wrong driver behaviour. Main reasons are deficiencies in adequate assessment of the actual traffic situation.

For these reasons the University of the Armed Forces, Institute of System Dynamics and Flight Mechanics has developed an intelligent **D**river **A**ssIsting **SY**stem (**DAISY**) which helps the driver to understand the actual traffic situation by hints or warnings when needed. However existing driver support systems suffer from low acceptance ratings by drivers. These systems usually don't take into account that different individual drivers might behave very differently when exposed to the same traffic situation. They do not adapt the warning messages to these individual driving traits. To overcome this shortcoming a driver support system has to be adaptive to both the actual situation and the individual driver.

1.2 DAISY components:

To satisfy these requirements a model based approach was used to implement the following main components of DAISY:

To account for the required situation adaptivity DAISY contains the ***Situation Analysis Module*** (SA). This module performs an explicit analysis of the current objective traffic situation by classifying it from a driver's point of view on the basis of environmental data. The result is a symbolic description of the situation. This information is used to predict the average behaviour of the other vehicles and to determine objective action limits caused eg. by other vehicles. The SA module is implemented on the basis of ***Petri-Nets***. This methodology allows a modular implementation covering different aspects of the total traffic situation and combining them in a very flexible manner. Examples of implemented situation models are car following and lane keeping /overtaking.

The ***Model of the Actual Driver*** (MOAD) accounts for the adaptivity to the individual driver. Here the actual driver's normal driving style is extracted from observations of his behaviour during a learning phase. In the MOAD the driver objectives (e.g. desired driving speed) are determined first. Then the driver intent is recognized by use of a rule based model. In a trained condition the skill based driver control inputs can be predicted and the individual normal driving behaviour described by characteristic time reserve values can be extracted. A further processing step is the assessment of the current driver resources.

The ***Warning System*** (WS) combines the output of SA and MOAD to issue a warning in the case of a dangerous deviation from normal behaviour caused for example by distraction, degraded vigilance or driver overload. To detect the deviation from normal behaviour the time reserve is used as a danger measure. The time reserve reflects the driving behaviour of danger prevention on the time basis by a combination of certain parameters. It not only takes into account the actual situation but also includes potential driver actions and potential actions of other vehicles which could endanger the own one. The time reserve proved adequate in a number of experiments (Kopf,1994).

Figure 1 shows the whole DAISY system. In the following section the design of the behavioural driver model as an essential component of MOAD is described in more detail; the skill based modelling is emphasized.

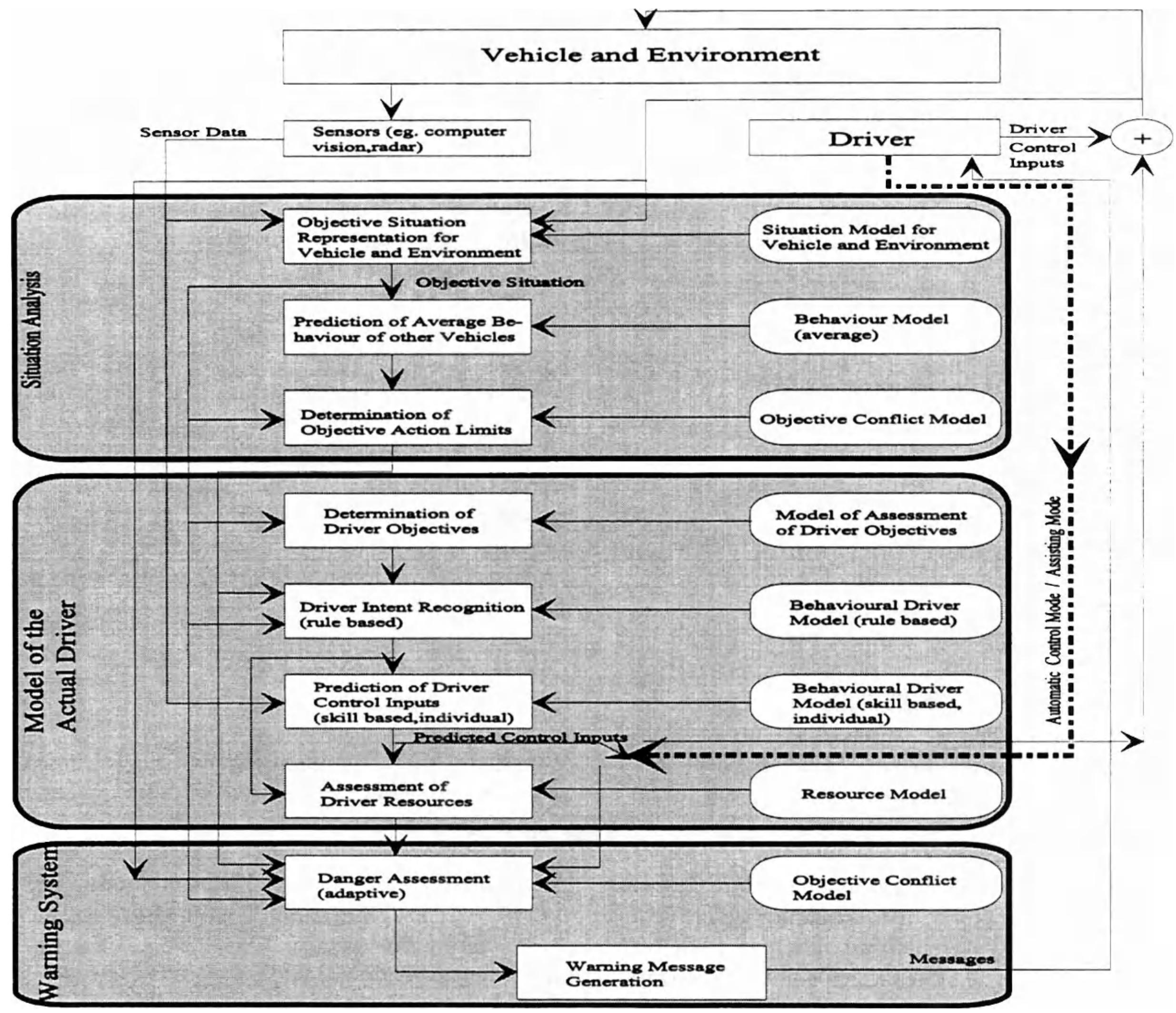


Fig. 1. DAISY architecture

2. DESIGN OF THE BEHAVIOURAL DRIVER MODEL

2.1 *Description of the problem*

Human behaviour can be classified in three main levels: Knowledge based, rule based and skill based behaviour (Rassmussen,1983). The driving task can be subdivided in the levels navigation, guidance and control which match the Rassmussen scheme in the following manner: Knowledge based behaviour - Navigation; Rule based behaviour - Guidance; Skill based behaviour - Control (Donges, 1992). DAISY mainly covers the rule based and skill based behaviour of the driver. It is one of the main design tasks to give the MOAD module the ability to learn the individual driving behaviour on the rule and skill based level.

When selecting an approach for driver modeling one has to pay attention to the following characteristics of the driving task:

- The driving behaviour is extremely situation specific.
- The driver perceives a lot of information at about the same time which influences his behaviour.
- Driving behaviour is nonlinear.

These characteristics can be met by an approach based on artificial neural nets. To select a neural net paradigm for driver modeling also technical requirements have to be taken into account. Necessary requirements are:

- The learning module must have realtime processing capability.
- Gathered knowledge about the driving behaviour must be stored (stability).
- New knowledge must be incorporated in the system (plasticity).

To meet also these requirements Fuzzy ART was chosen for the modeling task.

2.2 *Methodology*

Basic ideas in the neural network literature: The following basic mechanisms which are useful for driver modeling are implemented in different paradigms of artificial neural networks:

- Self organization
- Stability - plasticity trade off
- Association

All ART networks were designed to operate in an environment that requires realtime capability by using self organization as a basic feature. They also incorporate the requirement of sufficient stability of learned patterns while always being able to acquire new knowledge (stability - plasticity trade off). In particular, Fuzzy ART allows to combine different analog sensor inputs to one feature vector and can be trained in realtime by a fast-commit, slow-recode option (Carpenter, 1991b). But Fuzzy ART primarily only performs classification of analog patterns. To build a driver model an association between the actual traffic situation and the anticipated driver reaction has to be made. In Carpenter (1991a) ARTMAP was published which consists of two ART-networks combined by an inter-art associative memory. Basically this topology combines two different feature vectors by association. An association mechanism is also implemented in the counterpropagation network (Hecht-Nielsen,1988) and other network paradigms.

Adaptation of Fuzzy ART to the problem: The first adaptation step was simply to associate time series of driver reactions with classes of traffic situations. Traffic situations are represented as feature vectors which include sensoric measurements of the vehicle itself, street parameters and information about other vehicles. These vectors were complement coded.
Driver reaction was sampled and the signals were combined to vectors of time series of entire reaction sequences. To model the driving behaviour for lateral control the steering wheel rate and for longitudinal control the break and gas pedal position were taken. After classification of the situation describing vector by the selforganizing mechanisms of Fuzzy ART a vector containing sampled driver reactions is associated with the detected class during the learning phase. Refinement of this vector is done with the same learning rule which is part of the ART networks (Carpenter, 1991b). With this adaptation the driving behaviour can be learned by observation; this principle was used in a similar way but with a different neural network paradigm in Kraiss (1992).

The second adaptation step was to build a hierarchy of Fuzzy ART networks. The reason for this step was the realtime requirement but also to provide better model accuracy. The Fuzzy ART networks of the first level perform a coarse clustering process. For every class found for the first level a second clustering process with higher resolution is made. Every class found for the second stage is associated with the driver reaction sequences.

The third adaptation step was driven by the idea that different elements of the feature vector will have varying relevance in different traffic situations. Therefore a weighting of each element must be done. This was accomplished by introducing relevance parameters in the Fuzzy ART algorithm. In particular the category choice function (Carpenter, 1991b) was expanded with the component depending factors k_i (equation 1). These factors can be looked at as extra degrees of freedom compared to the basic algorithm which allows a situation specific weighting of the components of the feature vectors.

$$T_j = \frac{\sum_{i=1}^{M} \left| k_i \cdot \min(I_i, w_{ij}) \right|}{\alpha + \sum_{i=1}^{M} \left| k_i \cdot w_{ij} \right|} \quad (1)$$

T_j ≡ matching value for category j
$\mathbf{w}_j$ ≡ M-dimensional weight vector of category j
$\mathbf{I}$ ≡ M-dimensional inputvector
$\mathbf{k}$ ≡ M-dimensional relevance parameter vector
α ≡ choice parameter

Figure 2 outlines the Fuzzy ART hierarchy how it is implemented in the MOAD module and it's interaction with the SA. The symbolic description of the current objective situation is processed in a selection module which corresponds with the rule based Behavioural Driver Model to choose the ART-network covering the situation. Depending on the situation the input vector components and the ART-parameters are selected. All these informations are used in the Fuzzy ART hierarchy. In the trained condition driver specific outputs like time reserve values and control signals can be recalled. For driver adaptive warnings the actual time reserve values and the normal ones are compared in the WS. In DAISY right now 17 different Fuzzy ART hierarchies for longitudinal control are implemented; typical examples are car following or car approaching. For lateral control there are 24 Fuzzy ART hierarchies; typical examples are lane keeping and overtaking.

Optimization of network parameters: Designing an artificial neural network to solve a specific problem often means the selection of network parameters by trial and error. Right now, no analytical method is available which provides problem specific parameter values of a special neural network. This is also true for the above described Fuzzy ART approach for driver modeling. In this approach the two vigilance parameters for coarse and fine clustering but also the introduced parameters for situation specific weighting of the input vector components must be tuned to get accurate driver models.

Genetic algorithms have proven to be a very flexible and generally useable method to solve search and optimization problems (Goldberg, 1989). They realize the trial and error process to find a good problem solution. This method was used to optimize the Fuzzy ART parameters off-line in the following way (Figure 3): Recorded data from test trials which describe the driving behaviour of a particular driver

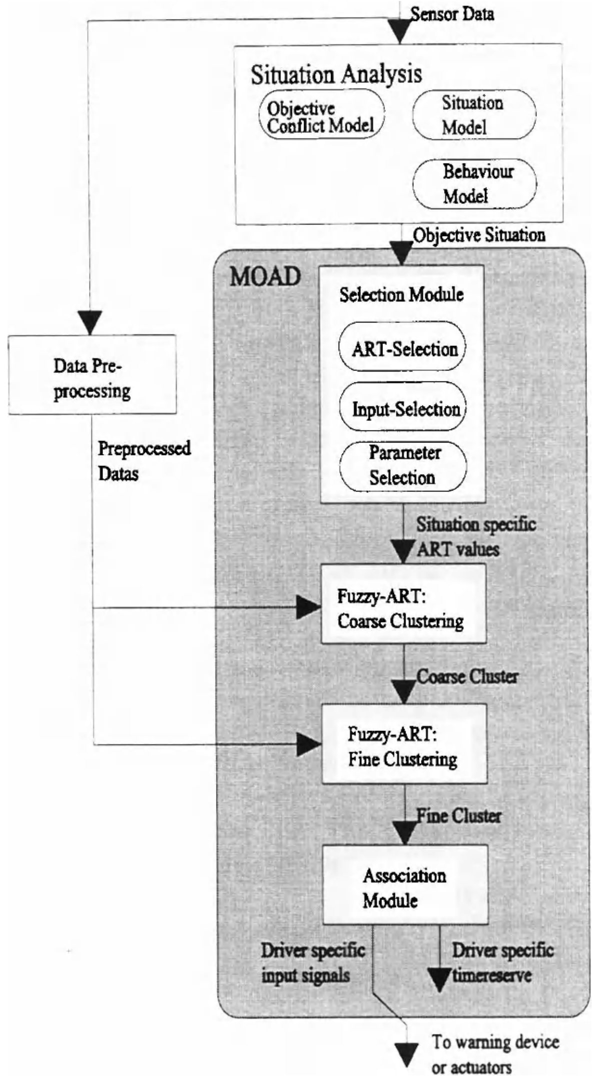


Fig. 2. The Fuzzy ART based MOAD module

are used to train a population of Fuzzy ART networks; each of these has it's own set of parameters (e.g. vigilance parameters or relevance parameters). After training every network is used as a controller to drive a model of the used car. During this closed loop simulation car reactions and previously recorded data of the human driver are compared in order to calculate fitness values for the different network controllers. These values are used in the genetic algorithm to adapt the network specific parameters by use of genetic operators. A binary representation of the Fuzzy ART parameters is transformed in the floating point values of these parameters which are used for a new training cycle. This process is repeated until a satisfying solution is found or a predefined number of generations is reached. After a parameter set is found this can be used to train drivers on-line in future. The procedure was used to first adapt the vigilance parameter set and then the relevance parameter set; it was done twice to stabilize the found parameters.

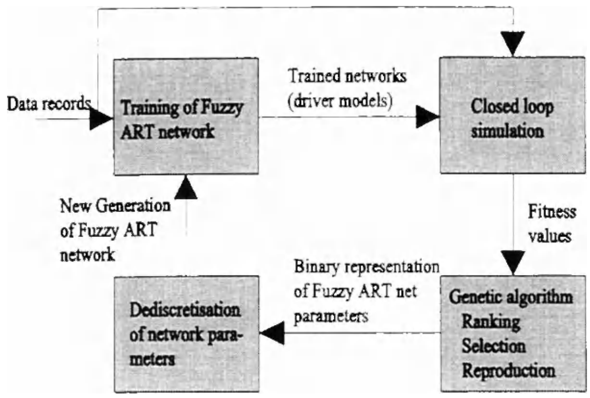


Fig. 3. Principle of network parameter optimization

The main characteristics of the implemented genetic algorithm are:

Steady state reproduction without duplicates
8 bit binary representation using gray coding
Linear normalization as fitness technique
Roulette wheel selection of individuals and operators
Used operators: One-point crossover, two-point crossover, uniform crossover,one-bit mutation
Population size: 16
New children: 4

The concept of Pareto optimality was used because different fitness values were evaluated at the same time to describe the total fitness of the individuals (Goldberg, 1989, pp.197 - 201).

2.3 Results:

First the use of a trained and optimized neural driver model as a human-like controller is shown for longitudinal control.Figure 4 shows a comparison of the driver and his trained network; the driven speed in a car following situation is depicted for both in a typical sequence. In Figure 5 distance keeping of driver and network in the same sequence is plotted. The network is stable and behaves very similar to the driver which is an indication that the model can imitate the driver.

As a sufficiently accurate human driver model is available this can be used in DAISY in different ways. Here the use for driver adaptive warnings for longitudinal control is outlined. The time reserve concept is well suited to define the warning threshold (Kopf, 1994). This measure can be used for objective as well as subjectively adapted hints and warnings. To realize situation adaptivity for each Fuzzy ART class class-specific statistical values can be evaluated for the time reserve eg. the average value and standard deviation. Both values describe the situation-specific normal region of the driving behaviour. By comparing the actual time reserve value with the normal region, deviations from normal can be detected. These deviations are considered to be caused by unnormal driving behaviour which may have it's deeper cause, for example, in a lack of vigilance. If the actual value of the time reserve is falling below a certain level adapted to the normal driving behaviour a warning is issued.

Figure 6 shows the actual time reserve values during normal driving compared to the situation specific threshold values of the normal driving behaviour region; as can be seen the driver does not leave the normal region during the entire sequence indicating that the driver behaves normally. Figure 7 shows the comparison of an imitated unnormal driving behaviour of the same driver with the learned normal behaviour region; in this case, the driver leaves his

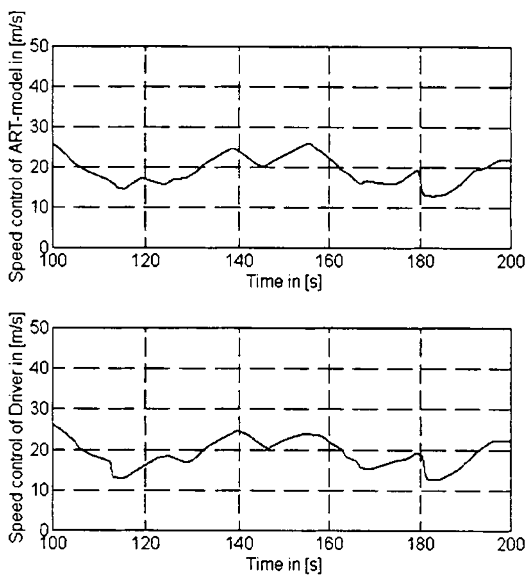


Fig.4. Speed control: Driver - ART model

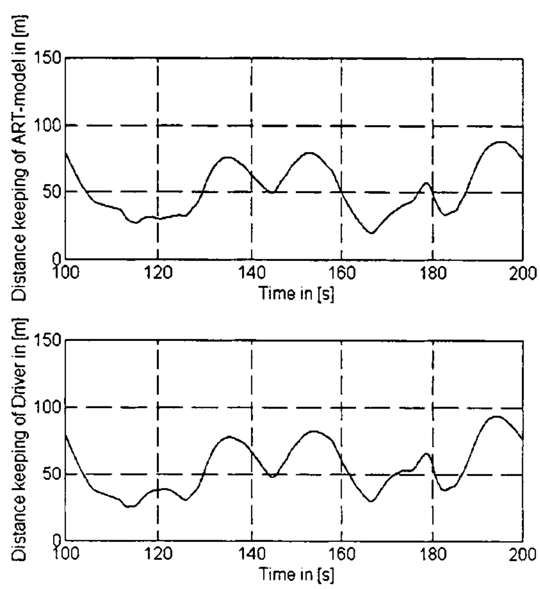


Fig.5. Distance keeping: Driver - ART model

normal time reserve region all the time; this deviation can be taken as an indication for unnormal driving. Some kind of message should be issued to the driver to draw his attention to this behavioural change. The intensity of this message should be increased along with further departure of the normal driving behaviour, in particular with respect to the case, when the time reserve is approaching zero.

3. CONCLUSION

In this paper it was shown that realtime learning of driving behaviour can be accomplished by a neural net approach based on Fuzzy ART. ART specific parameters and the introduced relevance parameters were found by means of off-line optimization with Genetic algorithms; once predefined, these parameter sets can be implemented in the on-line learning process to extract the normal driving behaviour by observation. The trained driver model can be used in a driver assisting system like DAISY in different ways: To satisfy the requirement of adaptiveness for messages and warnings it is possible to extract the driver and situation specific time reserve values to describe normal driving behaviour on a time basis. Deviations from these regions can be detected and used to issue driver adaptive warnings. It is also possible to use a trained neural network as a human-like controller. Other possible applications like driver intent recognition, assessment of current driver resources and driver identification are part of the current reseach work.

REFERENCES

Carpenter, G.A., S. Grossberg and J.H. Reynolds (1991a). ARTMAP: Supervised Real-Time Learning and Classification of Nonstationary Data by a Self-Organizing Neural Network.Neural Networks, Vol. 4, pp. 565 - 588

Carpenter, G.A., S. Grossberg and D.B.Rosen(1991b). Fuzzy ART: Fast Stable Learning and Categorization of Analog Patterns by an Adaptive Resonance System. Neural Networks, Vol. 4, pp. 759 - 771

Donges, E. (1992). Das Prinzip der Vorhersehbarkeit als Auslegungsprinzip für Maßnahmen zur Aktiven Sicherheit im Straßenverkehrssystem.VDI-Gesellschaft Fahrzeugtechnik, 1. Fachtagung

Goldberg, D.E. (1989). Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley Publishing Company

Hecht-Nielsen,R.(1988). Applications of Counterpropagation Networks. Neural Networks, Vol. 1, pp. 131 - 139

Kraiss, K.F., H. Küttelwesch (1992). Identification and Application of Neural Operator Models in a Car Driving Situation. IFAC Man-Machine Symposium 1992. Preprints pp.2.2.1

Kopf,M. (1994). Ein Beitrag zur modellbasierten, adaptiven Fahrerunterstützung für das Fahren auf deutschen Autobahnen. Doctoral thesis. VDI Fortschrittsberichte, 12, Nr. 203

Rassmussen, J. (1983). Skills, Rules, and Knowledge: Signals, Signs, and Symbols, and Other Distinctions in Human Performance Models. IEEE Transactions on Systems, Man, and Cybernetics,Vol. SMC-13, pp. 257 - 266

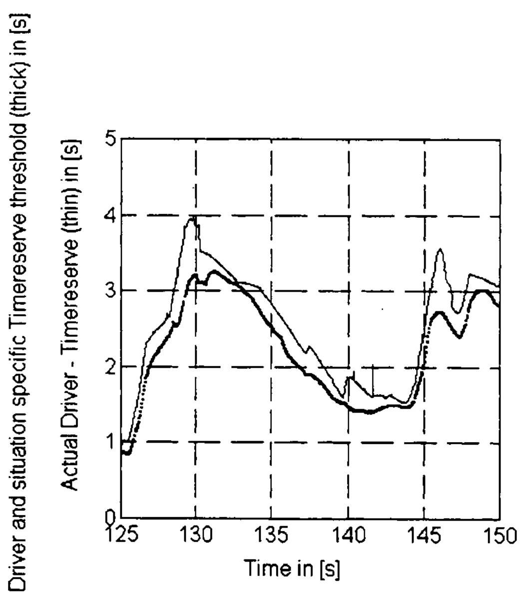


Fig. 6. Comparison of situation specific actual time reserve values during normal driving with the thresholds of normal behaviour region

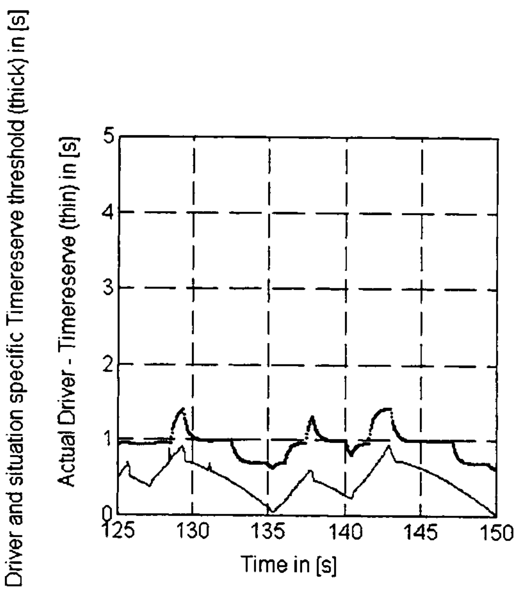


Fig. 7. Comparison of situation specific actual time reserve values during unnormal driving with the thresholds of normal behaviour region

CAR-FOLLOWING MEASUREMENTS, SIMULATIONS, AND A PROPOSED PROCEDURE FOR EVALUATING SAFETY

S. Chen*, T. B. Sheridan* , H. Kusunoki and N. Komoda****

**Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

***R&D Planning Division, Toyota Motor Corporation, Tokyo, Japan*

Abstract. Video data of highway traffic were used to study the effect of illumination, weather, and traffic density on the driver car-following behavior. A driving simulator was used to evaluate the effect of driver intention, gender, and experience on braking reaction time and following distance. The results show that the environmental factors influenence driver behavior mainly in congested traffic but not in free-flow traffic. Driver experience, gender and intention to pass all show some effects on car-following behavior. The distribution of driver reaction time fits a log-normal model. A safety evaluation procedure is proposed based on the obtained statistical distributions.

Key Words. Driver behavior; simulators; simulation; statistical analysis; safety analysis.

1. INTRODUCTION

Rear-end collision is one of the most common types of crashes involving two or more vehicles. The National Safety Council (NSC) reported (Accident Facts, 1991) that there were approximately 11.5 million automobile crashes in 1990, of which 2.2 million were front-to-rear-end, about 19.1% of the total. These crashes accounted for 24% of all collisions involving two or more vehicles. The NSC further reported, in the same year, on 1800 fatalities (in front-to-rear-end crashes) or 4.4% of the fatalities occurring in all motor vehicle crashes. Front-to-rear-end collisions accounted for 11.4% of all fatalities involving two or more vehicles in crashes (McGeheen et al, 1992).

There is plenty of evidence that these front-to-rear-end accidents result primarily from the following- drivers keeping inadequate headway for the speed of travel and, therefore, not being able to stop or slow down sufficiently when the lead vehicle unexpectedly decelerates or stops rapidly due to the late timing of a maneuver, an obstacle, or some other emergency situation. The NSC (Accident Facts, 1991) reported that 8.7% of all collisions or nearly one-half of front-to-rear-end crashes are due to drivers following too closely. Various warning systems and intelligent cruise controllers are being designed to help the inattentive, inexperienced or risk-prone driver keep an adequate following distance and react when unexpected situations occur. To be acceptable to the driver, those devices should be actuated by a "smart" algorithm which activates either or both warning and deceleration, depending on the circumstances.

What drivers believe to be an acceptable following distance is not easy to deal with for many reasons. The individual driver has his/her own intuitive and heuristic internal model of the following situation. The driver may also decide to engage in active risk-taking behavior when planning and carrying out driving maneuvers. The driver's sensory and perceptual limits for detecting speed and distances may change according to age and gender. These and many other factors are not mutually exclusive, which makes the defining of a safe following distance very complex.

Considerable research has been directed to drivers' car-following behavior or headway distribution. Tolle (1976) compared different models of headway distribution using real highway data. Koshi et al (1983) suggested possible dual-mode car-following behavior. Chishaki and Tamura (1984) formulated a headway distribution model based on the distinction between leaders and followers. Several researchers tried to define so-called "safe" following distance (or headway). Dull and Peters (1978) defined the safe distance as a function of vehicle braking capability and of driver reaction time. Fenton (1979) proposed a headway safety policy for automated highway operations. Ioannou et al (1992) used the California rule for safe car-following distance, a vehicle length for every 10 m.p.h, for their intelligent cruise controller. These mathematical models, though easy to implement in simulation, can hardly reflect the real driver judgments of the safe following distances, since they all try to define what is a safe following distance instead of what following distances drivers believe to be safe and/or actually employ.

This paper first investigates constant velocity car following behavior as a function of absolute speed, weather, illumination, traffic density, driver intention, experience and gender. Two approaches were used to measure the car-following behavior: field measurements and simulator tests. The field measurements provide realistic but uncontrolled data for a large and varied population under different environmental conditions. The simulator tests, on the other hand, provide data among different groups of drivers in a planned, controlled environment. Those data are then used to develop and validate a car-following evaluation procedure.

2. FIELD MEASUREMENTS

2.1. *Location*

Most measurements were carried out on Interstate I-93 near the downtown Boston area. This section of I-93 is a three-lane freeway going north and south. This location has the merit of having steady traffic flow during the rush hour and even during the non-rush hour. Only those data from the inner two lanes were taken. A high elevated parking lot along the highway was chosen to place a video camera so that we could have a clear view of the whole highway and thus were able to distinguish between vehicles on adjacent lanes. Another location, the Massachusetts Turnpike (I-90), was used to collect high speed data and to show the effect of different locations. This section of I-90 is a four-lane toll highway going east and west. Again, only the inner two lanes were considered. Both highways are well maintained and have good surface quality. The average throughput ranges from 1500 to 2000 vehicles per hour per lane during rush hour and approximately 1000 to 1600 vehicles per hour per lane during non-rush hour.

2.2. *Data Collection and Analysis*

A Sony Hi-8 Handycam video camcorder was used for measurement. It had a 10:1 zoom lens with 1 lux light sensitivity. The camera was adjusted so that about three cars following one another could be seen on the screen. This ensured proper car-following action but did not sacrifice accuracy. The camera had a built-in clock with resolution of 1 second. The traffic data collected from the highways were analyzed using the stop-frame method on a VCR. The VCR had a slow-frame speed of 1/60 second per frame, and this was used to calculate a velocity for the vehicles.

Since this study only considered car-following under a steady traffic stream, cars with acceleration/deceleration were discarded. To achieve this goal without measuring the acceleration rate, we employed the data only when there were at least three cars following each other at approximately the same speed. Two stationary marks, 55 feet apart on the highway, were used to calibrate between the distance that appears on the TV display (TV-screen-inches) and the distance on location (real distance). In order to avoid errors caused by image distortion from the camera, we measured the distance from the center of the screen where the markers were observed.

Each cited instance of car-following was classified by the speed at which both cars were traveling and the distance the cars were measured to be apart from each other. It was assumed and afterward checked that both cars were traveling at the same speed. It was also assumed that the velocity was constant for the specified distance between the two marks. A collection of about 50-100 of these data points constituted one data set for each video session. There were 60 video sessions from I-93 and 15 sessions from I-90.

2.3. *Results*

Based on the collected data, the car-following distance was found to vary greatly in each speed range. By observation, each distribution resembles a log-normal distribution.

Box plots are used for graphic comparison of the means of samples under different conditions. A box plot has several graphic elements. The lower and upper lines of each "box" are the 25th and 75th percentiles of the distribution. The line in the middle of the box is the median value of the samples. The lines extending above and below the box show the total range of data, except that any sample value that is more than 1.5 times the box range is shown as individual point above or below the box. Figure 1 compares the mean following distance under dry and wet road surfaces. An analysis of variance (ANOVA) of the effect of wet surface shows that there were no statistically significant differences for speeds higher than 40 km/hr, while there were significant differences for speeds lower than 40 km/hr ($P<0.05$). This suggests that a wet surface affects car-following behavior for congested flows, but not for free flows.

The same phenomenon exists for night time (Figure 2). An ANOVA shows that, except for the speed range of 30-40 km/hr, there were significant differences in the car-following behavior under congested flows ($P<0.05$) while it shows no significant differences under free flows. For the effect of traffic density, Figure 3 and its ANOVA seem to suggest that rush hour traffic has no effect on car-following behavior compared to non rush hour.

Overall, the environmental effect of pavement wetness and spacious illumination on driver car-following behavior seems to be minimal when the traffic is in free flow, while the effect appears to increase when the traffic turns into congested flow. As for traffic density, no obvious effect is shown for the collected data.

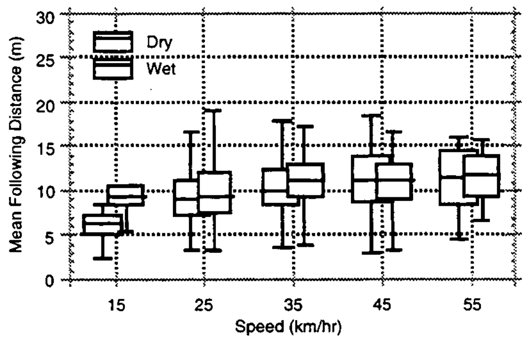


Fig. 1. Comparison of Mean Following Distance under Dry and Wet Road Surface

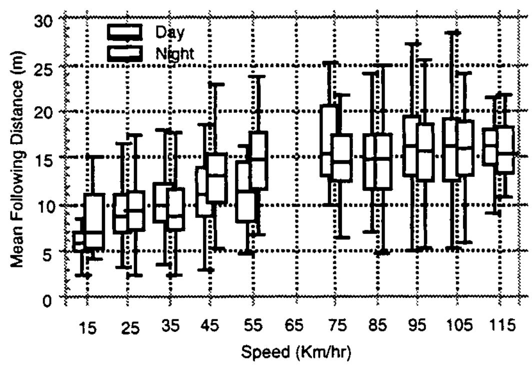


Fig. 2. Comparison of Mean Following Distance under Day and Night

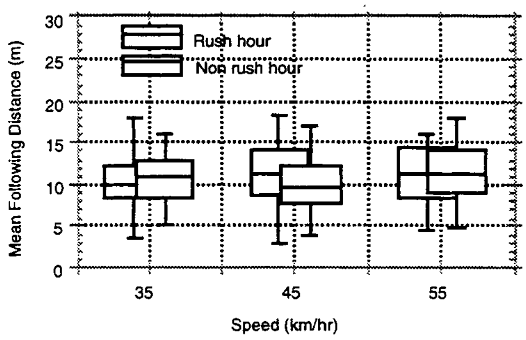


Fig. 3. Comparison of Mean Following Distance during Rush and Non-rush Hour

3. SIMULATOR TEST

3.1. *Human-Machine Systems Laboratory (HMSL) Driving Simulator*

In most cases, the rationale for developing a driving simulator is to provide a safe and economical means for presenting an operational scenario in a controlled environment with readily available measures of system performance. Many examples of the acceptance and utilization of driving research simulators now exist throughout the United States and Europe.

The driving simulator developed in the Human-Machine Systems Laboratory provides a variety of possible parameter variations. Figure 4 shows the configuration of the simulator. The fixed-base car cabin allows the driver to control the vehicle through the existing gas pedal, brake, and steering wheel. Two audio speakers are used to generate the sound environment. Steering wheel torque feedback is generated by a computer-controlled DC motor on the steering shaft. The motion of the vehicle is fed back to the driver mainly by the visual cues from the projection screen and secondarily by the feel of the steering wheel and the auditory cues which include engine noise and a car-passing sound. Both sound effects change in loudness and pitch according to the closing speed and distance relative to on-coming cars. The road scene consists of a two-lane highway with vehicles in each lane. The subject's vehicle can pass or be passed by other vehicles.

Generally speaking, skilled drivers were able to sense the lateral acceleration through visual cues and steering wheel torque. This simulator has been tested and performance on it compared to results from actual road test with regard to longitudinal distance to objects, oversteer on sharp curves and other variables.

3.2. *Experimental Task*

The simulator was used to test the effect of driver experience, gender and intention to pass on car-following behavior, and to test the driver's braking reaction time under different conditions, such as lead car applying full brake and a suddenly appearing obstacle. In other words, car-following

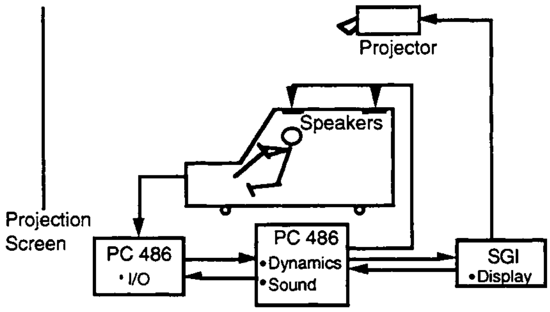


Fig. 4. Configuration of HMSL Driving Simulator

distance distribution and reaction time distribution under different driving conditions and under different speeds were measured using the HMSL simulator.

The subjects were first given oral and written instruction describing the purpose and the procedure of the experiment. They were then asked to practice driving the simulator for as long as necessary in order to easily handle the vehicle and be familiar with a variety of presented scenarios that would appear in later experiments. Immediately following the practice, two different experiments were conducted with about 12 minutes for each experiment and a 5-minute break between them. In the first experiment, the subjects were instructed to follow the lead car and to change speed or to brake as they would in normal driving according to the lead car conditions. The lead car changed speed randomly and occasionally braked to stop. For the second experiment, the subjects were allowed to pass whenever possible without collision. There were "no passing" zones or on-coming cars to limit the passing frequency. Following distances at different speeds were continuously recorded in the computer, and braking reaction times were measured whenever the lead car applied its brake. This reaction time was measured from the onset of the lead car braking until the onset of the subject's braking.

3.3. *Subjects*

The 24 drivers tested in the simulator were all holders of a current U.S. driver license. None of the drivers was professional. The subjects had been selected on the basis of their driving experience and their gender. Subjects included:

(1) Inexperienced drivers, with a mean of 1.3 years of driving (σ = 0.5 year), and a mean age of 27.2 (σ = 2.2).
(2) Experienced drivers, with a mean of 6.3 years of driving (σ = 1.8 years), and a mean age of 29.7 (σ = 5.5).
(3) Female drivers, with a mean of 3.8 years of driving (σ = 2.4 years) and a mean age of 29 (σ =5.2).
(4) Male drivers, with a mean of 4.0 years of driving (σ = 3.1 years) and mean age of 28.3 (σ = 4.1).

Categorization into two experience levels was made on the basis of statistical risk of accident, assuming drivers with up to 2 years experience have a higher risk of accident (Colbourn et al, 1978).

All subjects were volunteers and were not paid for their services during the period of approximately 1 hour required for testing.

3.4. *Results*

The log-normal distribution was assumed to compare the mean following distance for each experimental condition.

The main effect of driver experience on following distances was statistically significant ($P<0.05$) for all speed ranges except for the range of 60-70 km/h ($P=0.34$). The comparison of the mean distance is shown in Figure 5. Figures 6 and 7 show the main effects of gender and intention to pass. Again, the differences are statistically significant in most speed ranges for both factors ($P<0.05$).

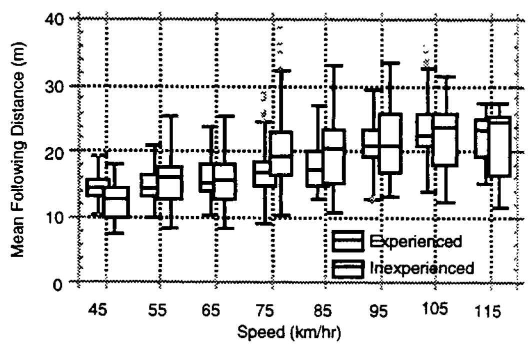


Fig. 5. Comparison of Mean Following Distance for Experienced and Inexperienced Drivers

Figure 8 shows the cumulative density function for braking reaction time. The mean reaction time for the lead car applying its emergency brake is 1.25 seconds, while that for sudden appearance of an obstacle is 1.10 seconds. Both reaction times include both perception time and foot movement time.

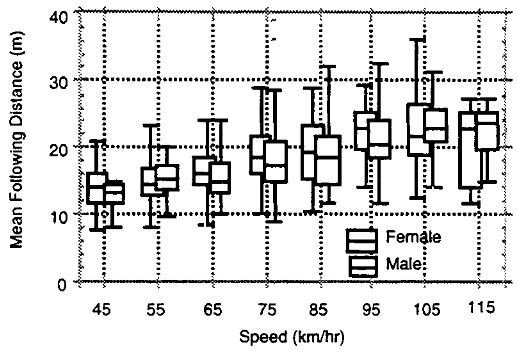


Fig. 6. Comparison of Mean Following Distance for Female and Male Drivers

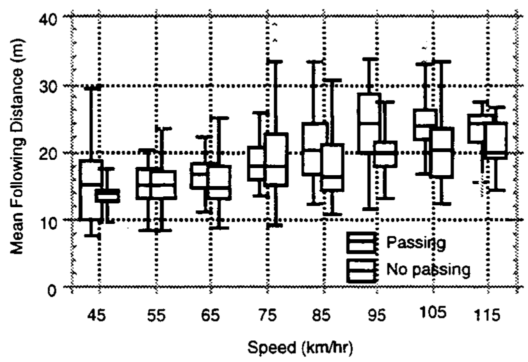


Fig. 7. Comparison of Mean Following Distance with Passing and No-passing

4. PROPOSED PROCEDURE FOR EVALUATING THE SAFETY OF CAR-FOLLOWING

Our analysis of the car-following strategy of a driver is based on all the factors considered in the previous two sections. Based on these data, the estimated probability distributions, a procedure is proposed to evaluate safety. It makes use of the commonly used equation for "safe following distance"

$$D = V_f T_r + \frac{1}{2}\left[\frac{V_l^2}{\alpha_l} - \frac{V_f^2}{\alpha_f}\right] \quad (1)$$

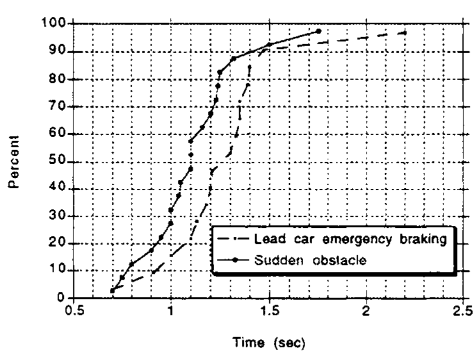


Fig. 8. Braking reaction time

where D = the stopping distance required to avoid collision, V_f = the speed of following car, T_r = the driver braking reaction time, V_l = the speed of lead car, α_l = the deceleration of lead car, α_f = the deceleration of following car.

The parameters α_f and α_l are dependent upon vehicle braking characteristics while T_r varies according to the probability distribution derived in section 3. For the extreme case of a sudden-stopped obstacle, α_l is zero and the required stopping distance would be the largest.

Figure 9 diagrams the proposed procedure for evaluating the safety of car-following. The car-following distance distribution is acquired using the data from both highway measurements and simulator tests. The distribution of driver reaction time, on the other hand, comes from the results of the simulator test only. The vehicle braking characteristics are dependent on vehicles type, vehicle speed, and road surface conditions. The discrete following distance and stopping distance are sampled separately using the Monte Carlo method. The difference between these two distances then forms the distribution of safety margin under that particular condition. The status of the vehicle, collision or not, is then determined.

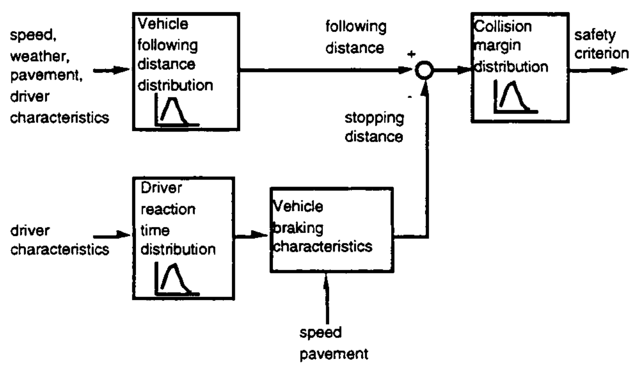


Fig. 9. Block diagram of the car-following evaluation algorithm

To illustrate the procedure, consider the following example. The distribution of the car-following distance under daylight, dry pavement, and for the speed range of 30-40 km/hr acquired from the highway data is shown in Figure 10. The distribution of the brake reaction time is described in Figure 8 for the the special case of encountering a sudden obstacle. In this case, it is further assumed that, in an emergency, the following driver will brake at or near limit of tire-pavement friction. Hence, the braking of the following vehicle is fixed at 0.7g, a level achievable on dry pavement by vehicles capable of meeting FMVSS 105, the Federal Motor Vehicle Safety Standard for braking. The stopping distance for this special case was derived by using Equation (1). This stopping distance distribution is also shown in Figure 10. Using the Monte Carlo sampling method, the safety margin could be a distribution as shown in Figure 11. From this hypothetical result one can see that under this special condition, about 96% of the population would have a collision if a sudden obstacle did appear. The same procedure can be used to evaluate the situation that the lead car applies full brake and slows to a stop. In that case only about 55% of events result in a collision.

The procedure can be modified with some correction factors for different assumed conditions but using similar statistical distributions. For example, data from the highway measurements show there is 0.9 meter greater average following distance under wet road surface for speed range 30-40 km/h. The safety margin might then be estimated by simply shifting the distribution of safety margin to the right by 0.9 meter. In this case the collision probability would be 94% instead of 96% as in the dry surface condition.

The procedure can be further refined to include correction at the stage of Monte Carlo sampling between reaction times and following distances (a young man with faster reaction times follows more closely) and of course can be improved upon with better data as it becomes available.

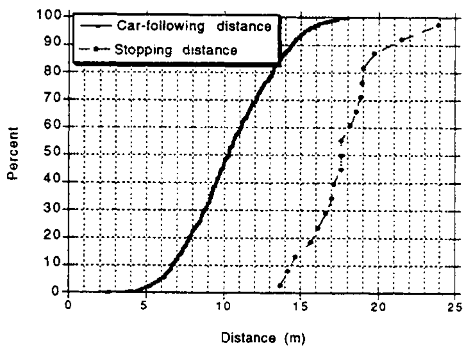


Fig. 10. Cumulative frequency distributions for car-following distance and stopping distance

5. CONCLUSIONS

A procedure for evaluating the safety margin of driver car-following behavior has been presented. Highway data on car following and simulator experiments on car following and braking were performed and the effects of different environmental conditions and driver conditions were discussed. The results show that there may not be a simple relationship between the car-following distance

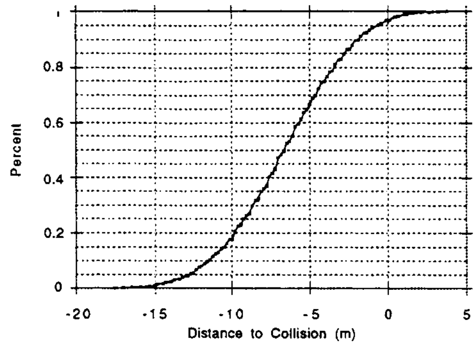


Fig. 11. Hypothetical distribution of the safety margin

and any single factor. This makes the study of driver car-following behavior difficult. A procedure based on these data plus assumptions about vehicle braking permits an evaluation of margin-of-safety. Further studies are needed to refine the procedure for potential application in the design of an intelligent cruise controller and other improvements in highway transportation.

6. REFERENCES

McGehee, D.V., T.A. Dingus and A.D. Horowitz (1992). The potential value of a front-to-rear-end collision warning system based on factors of driver Behavior, visual perception and brake reaction. *Proc. of the Human Factors Society 36th Annual Meeting,* Atlanta, 1003-1005.

Chishaki, T and Y. Tamura (1984). Headway distribution model based on the distinction between leaders and followers. *Transportation and Traffic Theory,* 43-63.

Tolle, J. E. (1976). Vehicular headway distribution: testing and results. *Transportation Research Record,* 456, 56-64.

Koshi, M., M. Iwasaki and I. Ohkura (1983). Some findings and an overview on vehicular flow characteristics. *Proc. of the 8th International Symposium on Transportation and Traffic Flow Theory,* 1981 (edited by Hurdle V.F., Hauer E. and Stewart G.N.) University of Toronto Press, Toronto, Ontario, 403-426.

Ioannou, P.A., C.C. Chien and J. Hauser (1992). Autonomous intelligent cruise control. *Proc. of the IVHS America 1992 Annual Meeting,*, Newport Beach, California, 1, 97-105.

Dull, E.H. and H.J. Peters (1978). Collision avoidance system for automobiles. *Expert System and Fuzzy Systems.* SAE paper 780163.

Colbourn, C.J., I.V. Brown and A.K. Copeman (1978). Drivers' judgements of safe distances in vehicle following. *Human Factors,* 20(1), 1-11.

Fenton, R.E. (1979). A headway safety policy for automated highway operations. *IEEE Transaction on Vehicular Technology,* VT-28, 1, 22-28.

HUMAN-MACHINE ORGANIZATION STUDY THE CASE OF THE AIR TRAFFIC CONTROL

F. Vanderhaegen*

Laboratoire de Mecanique et d'Automatique Industrielles et Humaines, URA CNRS 1775
Université de Valenciennes et du Hainaut-Cambrésis
B.P. 311 - Le Mont Houy - 59304 VALENCIENNES Cedex - FRANCE
E-mail: vanderhaegen@univ-valenciennes.fr

* : at present, Post-doctorate in the Joint Research Centre of the European Commission
TP210 - ISEI - 21020 Ispra (VA) - ITALY
E-mail: frederic.vanderhaegen@cen.jrc.it

Abstract: This paper presents human-machine organization study based on the working position concept. A methodology for designing this kind of system is proposed. The resulting organization is then described as a complex communication network between both the human and technical control systems of each working position. This global representation is illustrated with the air traffic control organization. Finally, the impact of new control tools is discussed through experimental and field studies.

Keywords: Adaptation, Air traffic control, Automation, Evaluation, Human factors, Multilevel system, System design, Work organization.

1. INTRODUCTION

In relation to the study of human-machine systems, an entire automated command is sometimes undesirable. Indeed, the corresponding automated system might not be entirely flexible nor reliable, in the case of a controlled process malfunctioning, for instance. Anyway, the long-term human inactivity in the control and supervisory loop of a complex system may lead to a loss of expertise. Moreover, an entirely manual command may also lead to some problems when for example, the human operator is overloaded. Therefore, when choosing the degree of automation for integration into a human-machine system, the main compromises include (1) what, (2) why and (3) how tasks have to be automated, according to both collective and individual features. In this way, the application of human-machine cooperation principles is a good way to avoid excessive automation, and to provide to the human operators specific cooperative assistance tools for decision and/or action.

The air traffic control domain confronts the problem of both the design and the validation of a future system organization. Indeed, (1) as the air traffic increases considerably, (2) as the human air traffic controller's workload becomes sometimes heavy, (3) as the present control tools are saturated, it is urgent to assist efficiently the air traffic controllers in order to optimize their activity regulation. In this way, the purpose of many studies is to propose and validate a new air traffic control organization which allows the air traffic controllers to stay active in the control and supervisory loop of the process and which permits maintenance of an optimal global system performance.

In such a way, the first part of this paper deals with this kind of human-machine system description. The second part evolves different organizations for the command system of a given controlled process. Finally, this global discussion is illustrated by the last part which concerns the air traffic control domain and in which concrete examples will give some practical system validation problems.

2. HUMAN-MACHINE SYSTEM MODELING

A human-machine system design can be oriented toward a global methodology which is helpful to describe the studied organization. This kind of system study is based on specific working position modeling in which communication can be made through different interaction tools.

2.1. System design methodology.

This methodology is directed toward a top-down design phase and toward a bottom-up evaluation phase, figure 1. Each phase makes use of two models: the human model which determines the human limits and resources, and the task model which defines the interactions and the information needs of the team in order to perform the individual and collective tasks.

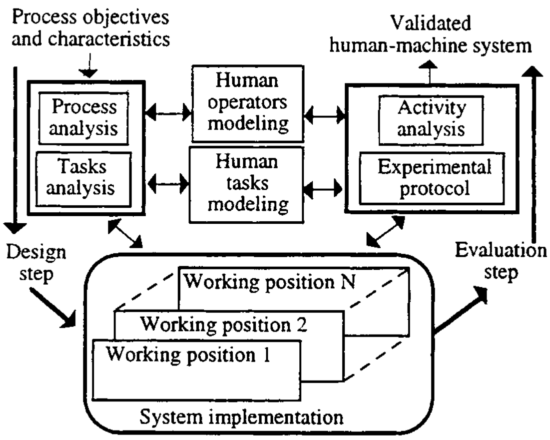


Fig. 1. System design methodology.

The top-down phase begins with a process analysis that permits the definition of the control and supervisory tasks. Prescribed tasks are defined depending on material and human resources constraints so as to obtain (1) the foreseen functioning modes to determine what information is needed for process supervision, and (2) the foreseen dysfunctioning modes to provide the actions sequences carried out by the human operators to solve these problems. The task analysis concerns tasks performed by each human operator and the interactions between the different members, e.g. human actions chronology, human communications. This analysis is based on both objective and subjective data collected either in real context during normal activities, or in simulated context during pre-defined experiments. It also determines the physical working position separation, the task sharing between the human operators and the legibility level of one human operator activity by another one. They lead to a task classification: those that may be entirely automated, those that cannot be executed by an automated tool, those that can be shared. In fact, in each working position, this classification allows the characterization of the human operator's role in the control and supervisory loop, the level of automation and the human-machine cooperation modes. Before the integration of the final human-machine system, a last step consists of the specifications of the dialogue interfaces.

The bottom-up phase, based on the study of effective tasks, aims at validating the human-machine system with regard to some evaluation criteria about individual activity and interactions between working positions. The evaluation objectives can take into account an assessment of the global performance, i.e. the difference between the real production of the piloted system and the expected production. Moreover, ergonomic criteria can explain some user problems, from methods that help to reconstitute the unobservable mental human activity. They are based on subjective evaluation methods and measurable parameters to assess for example the human workload. An appropriate experimental protocol defines the evaluation contexts and criteria to take into account during experiments. Finally, the analysis of the recorded subjective and/or objective data of the effective performed tasks will permit validation or refinement of the proposed human-machine organization.

2.2. Working position modeling.

In the basic working position architecture, human-machine interactions are numerous and occur through two control systems (Vanderhaegen, 1993), figure 2.

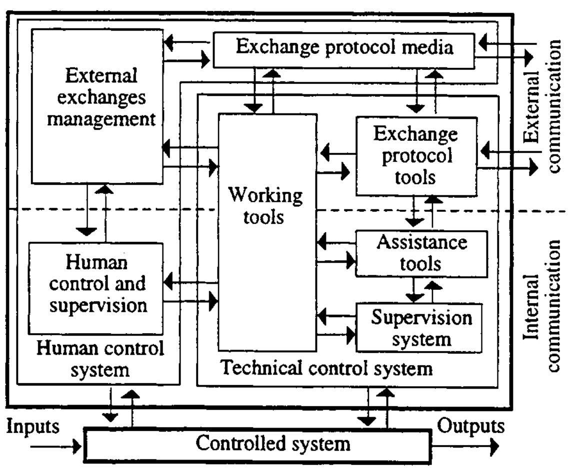


Fig. 2. Working position modeling.

The human control system is constituted by one or several human operators. The technical control system is constituted by a local supervision system, one or several assistance tools, working tools and exchange protocol tools. The supervision system manages the input and output data from the controlled process. On the other hand, the assistance tools provide the human control system with a help to action and/or decision. The human operators are integrated into the main control and the supervisory loop of the process. Using different working tools that provide them all the information they need, they manage the local control goals and the process abnormalities that the automated system (i.e. both the assistance tools and the supervision system) cannot solve. They also take into account the collective work goals. In such a way, they manage the data they have to send to and/or to receive from other working positions.

The representation of the human-machine system includes these external and internal data flows. Internal communications correspond to the specific interactions that exist inside a working position according to local work goals. External communications are the result of communication procedures that exchange data between different working positions using (1) human communication protocol media (e.g. verbal or gestual communication, manual transfer), (2) specific working tools (e.g. telephone, radio, screen, printer, pen, paper, office supplies) and/or (3) exchange protocol tools (network, telephones lines, aerial, radar). External human-human communications are established either by specific verbal or gestual communication protocol, or by manual transfer of working tools such as a paper support. Sometimes, they are made through specific working tools as telephone or radio. In such a case, they require additional connection supports like telephone lines and aerials, in order to make the external communication possible. This transfer data can also be generated automatically by the assistance tools.

3. COMMAND SYSTEM ORGANIZATION

One of the results of the general system design methodology is the definition of the human-machine cooperation which presents different possible connections between assistance tools and human control systems. In this context, cooperation includes the coordinated tasks processing, taking into account geographical processing area and actions chronology between decision-makers. In such a way, the command system needs specific interfaces and can be managed from different interactions between working positions.

3.1. Assistance tools connection.

Interactions between the human operators and the assistance tools through the working tools are multiple, figure 3. These interactions are related to the definition of assistance tool goals.

According to a vertical human-machine cooperation, assistance tools contribute to the decision-making process. In fact, human operators are responsible for the process tasks, and, if necessary, they can call on decision support tools which will supply them with advice.

Related to a horizontal human-machine cooperation, assistance tools contribute to the course of action, i.e. the on-line decision assistance tool can act directly within the process. The human control system and the assistance tools are on the same decision level in which the control and supervisory tasks, and their related actions can be distributed in order to relieve the human operators of overloaded situations. Thus, the dynamic task allocation principle consists of inserting in the control and supervisory loop, a task allocator that shares control tasks between decision-makers. This task allocator can be managed according to subjective or objective allocation criteria, by taking into account a global performance score, a mental workload measure or an assessment of task demands.

With regard to a multilevel human-machine cooperation, the assistance tools of a given working position help to filter the data between human and technical components of other working positions.

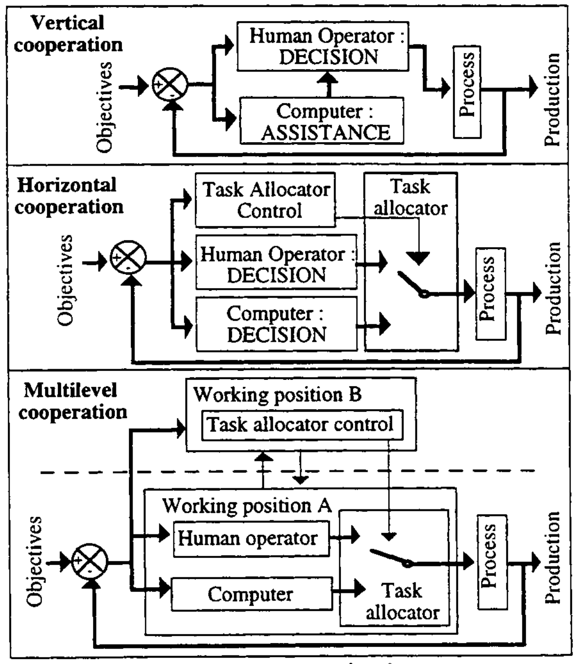
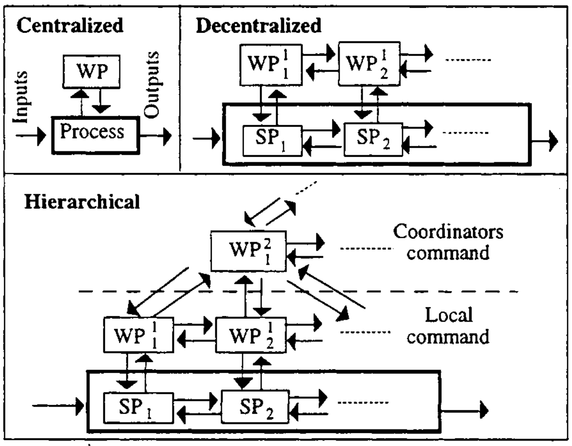


Fig. 3. Command system organization.

Those connections of the assistance tools require optimal communications between human operators and the technical control system. The set of working tools for those communications constitutes the human-machine interfaces. In a working position, different kinds of interface can be considered depending on the global work objectives. Nowadays, interface design is oriented toward highly-evolved systems which are more and more adapted to human operator behavior. For example, there are flexible personalized interfaces which adapt (automatically or manually) to modifications of the human environment. Thus, different ways of adapting the interfaces to the requirements of the task, the machine and the human operator can be considered. For instance, with regard to the human-machine cooperation level choice, interfaces can adapt (1) to the human operators' ability or their skill level, (2) to the user by changing conversational style, (3) to the concept of: "What you see is what you get", and/or (4) to a mental workload assessment. In such a way, adaptive and multimodal interface principle can be applied. It is based on interfaces with different kinds of dialogue so as to make intelligent working positions. As a matter of fact, working positions with integrated dataglove, touch screen, speech recognition, voice synthesis and video, can be studied. Furthermore, the concept of groupware, e.g. anthropocentric production systems, computer-support cooperative work systems, group decision support systems, can also be used. The general and common purposes of those systems are to facilitate the communications and the data sharing between the working positions components, in order to free them from restraints of time and/or space, to improve their productivity and the system safety.

3.2. Working position organization.

Therefore, a working position organization aims also at regulating the human activity. Indeed, the human and technical tools enable regulation both by retroaction and anticipation or by defining different filtering levels between working positions. In such a way, organizations can be structured so as to manage data sharing and make negotiation possible in case of a lack of information or knowledge. Depending on the connections between working positions, different kinds of human-machine organizations can be represented. Figure 4 gives examples of such a representation.



WP_i^j : Working Position i of the level j of the organization
SP_i : Sub-Process i

Fig. 4. Examples of human-machine organization.

The centralized organization is characterized by a single working position that takes into account the whole command process. In such a case, the media for external communication are not used. In other types of organizations in which the global process can be divided into different semi-dependant sub-processes, the decentralization enables the distribution of tasks over working positions. The control system components have to cooperate so as to synchronize, exchange data, or access a common resource. Nevertheless, in the case of a complex process, it seems necessary to coordinate each individual component so as to optimize the overall process. This function can be performed by other upper command levels, i.e. the coordinators in a hierarchical organization.

4. THE AIR TRAFFIC CONTROL CASE

Air traffic control is a good concrete example of a multilevel system which allows a control data filtering from different control levels. Particularly, the present control unit organization will change according to the variable abilities of human and technical control systems of the corresponding working positions. This entails the possible reorganization of the human operators' role in the control and supervisory loop. This point is developed through illustrative studies.

4.1. Global air traffic control organization.

The global organization of the Air Traffic Control can be represented as a hierarchical one, figure 5.

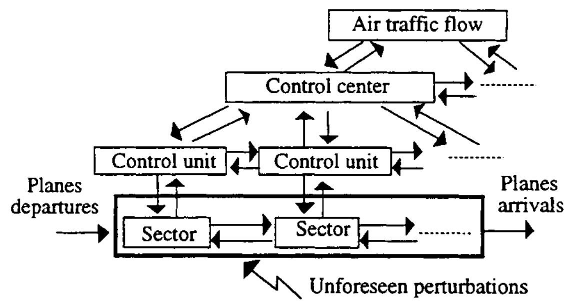
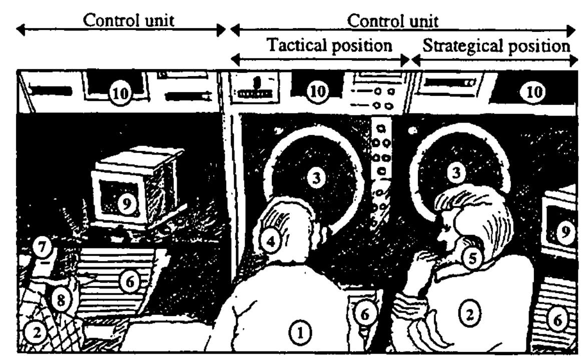


Fig. 5. Air traffic control organization.

The upper air traffic flow management is rather a long-term planning and regulation of the traffic flow, whereas the control center level concerns a medium-term management of the control units. Both levels organize the traffic (1) quantitatively because they act on the number of planes taken in charge by the lower control units level, and (2) strategically because they do not command the controlled process directly. A control unit supervises a given geographical sector. It is decomposed into different working positions managed by human controllers, such as a tactical radar controller and a strategical short-term planning controller, figure 6.



1 - Radar controller 2 - Planning controller 3 - Radar screen
4 - Radio 5 - Telephone 6 - Strip board
7 - Strip printer 8 - Paper strip 9 - Stripping display
10 - TV channel to display operational data (meteorology, military sectors)

Fig. 6. Example of control unit organization.

Through a radar screen, the radar controller supervises the traffic and dialogues with aircraft pilots by means of the radio. The supervision consists of anticipating possible conflicts between planes that may transgress separation norms and in solving them asking pilots to modify their initial trajectory. Using the telephone, the planning controller communicates with those of adjacent sectors in order to ensure the coordination between sectors and to avoid unsolvable conflicts at the sectors borders. He also anticipates the traffic density with regard to the sector capacity and regulates the radar controller's workload.

Both radar and planning controllers use paper strips, strip board to class them, radar screens that display on-line the evolution of planes and operational data that are displayed on a TV channel. Each printed strip contains the flight plan data of an aircraft. The planning controller receives it before the plane enters the sector. He analyses it, and transfers it to the radar controller who sorts it according to the dynamic planes evolution. Sometimes, from a simple stripping display, he can modify directly a flight plan. At this lower control unit level, the controllers evaluate qualitatively the situation because they optimize the control quality in terms of (1) safety constraint to avoid possible conflicts, (2) economic criterion to minimize the planes consumption and (3) regularity criterion to take off and land at the expected times.

4.2. Experimental case study.

This experimental study, called SPECTRA (French acronym for: Experimental System for the Air Traffic Control Tasks Allocation), is based on a multilevel cooperation concept (Vanderhaegen, 1993) , figure 7.

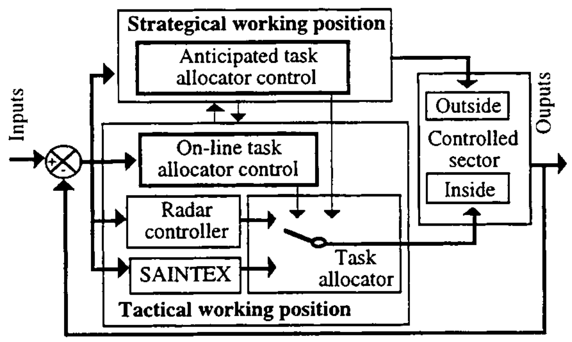


Fig. 7. Control unit structure on SPECTRA.

This study is carried out at the University of Valenciennes, in cooperation with the CENA (French Research Center in Air Traffic Control). It aims at increasing the sector capacity by regulating the controllers' activity and maintaining an optimal global performance. On the one hand, at the tactical level, the assistance tool is based on the dynamic task allocation principle that consists of sharing control tasks between a radar controller and an expert system, called SAINTEX (French acronym for: Experimental Automatic System for Night Air Traffic Control). On the other hand, at the strategic level, it concerns a system that cooperates with the planning controller and that is oriented toward a scheduling assistance tool for the optimization of the radar controller's activity regulation.

Firstly, a feasibility study of the horizontal cooperation was made at the tactical level (Vanderhaegen, *et al.*, 1994). In such a way, both a manual and automated dynamic allocation modes were integrated in SPECTRA structure. The manual one was performed by the radar controller who estimated his own performance and workload, and who allocated tasks either to himself or to SAINTEX. The automated one was based on the capabilities of the two decision-makers. For SAINTEX, those abilities are functional ones: only conflicts between two planes can be treated. For the human controller, these abilities are linked with functional task demands assessment which characterizes the task complexity. When task demands are too high and exceed a maximum level, conflicts are allocated to SAINTEX. In order to validate the SPECTRA concepts, four experiments (a training one, without assistance, in manual and automated modes) have been performed by nine qualified and six inexperienced controllers. The working tools were limited to a radar screen and a specific electronic stripping display to control planes by means of a keyboard and a mouse. Data related to (1) global performance related to the safety and economic levels, (2) task demands assessment, (3) subjective workload estimation on a unobtrusive graduated board, (4) subjective global workload with the Task Load indeX method, and (5) subjective comments about general information, workload information, comparisons between allocation modes, have been recorded during each experiment.

The main results are optimistic. Indeed, depending on obtained subjective and objective data, both manual and automated dynamic tasks allocation modes are useful for loaded traffic control, in terms of workload decreasing and safety improvement. On the one hand, qualified controllers tend usually to verify the SAINTEX strategy because they do not entirely trust it. Anyway, they preferred the manual mode which allowed them to feel fully responsible for the entire air traffic control. Nevertheless they performed best with the automated mode which obliged them to take in charge fewer conflicts. On the other hand, the inexperienced controllers preferred the automated mode which released them from the extra load of task sharing. They have the best result with the manual one because they have given to SAINTEX more conflicts. Anyway, despite different uses of SAINTEX abilities, both a manual and automated mode permits adaptation of the task sharing related to the human need for assistance.

4.3. Field case studies.

In France, a national progressive restructuring of control rooms is planned from 1995. In this way, a CENA project, called PHIDIAS (French acronym for Integrated Peripheral of Dialogue and Assistance) was created in order to design and validate possible new tools which might be integrated on the future control unit for the air traffic control (Salvi, 1992). This groupware validation is firstly based on a strong experimental study called HEGIAS (Host for Experimental Graphical Interfaces of Advanced automatic Systems). Through HEGIAS, some controllers from different control centers have validated the main human-machine interface principles proposed by PHIDIAS. Indeed, the intelligent electronic stripping that might replace the paper strip, the dialogue means (i.e. a touch screen, a mouse and a keyboard) and the codification by colour use seem to be helpful. This experimental step is convincing. Nevertheless, this new organization will not integrate anymore the manual strip management that is a good support for the human operational memory by means of physical writing and handling. On the other hand, an electronic stripping will permit the integration of assistance tools for decision and/or action, in the second step of the global PHIDIAS project. Therefore, one of the main work environment change is the transfer from an important human communication with verbal communications and paper strips exchanges to a human-machine-human communication by means of an electronic stripping which allows the user to receive and send information. In such a way, PHIDIAS might modify the work environment in which human controllers will work in a more virtual world. Therefore, despite a positive experimental step, the main problem is to know if PHIDIAS will be a soft transition or a hard rupture (Morico and Poirot-Delpeche, 1992).

Another interesting field study concerns the case of Linate airport, in Italy. Indeed, in April 1994, the control room work environment was modified. The transfer from the old control room to the new one was characterized by a space restructuring and control tool modifications. At the beginning, important economic consequences were due to lateness of planes. On the one hand, the lack of efficient training in the new control room and some imperfections of the new system entailed reducing temporarily the sectors capacity in order to minimize the safety risks. On the other hand, the expected number of planes was not decreased by the companies related to this modification. As a matter of fact, the lack of training in a simulated context entailed a hard change of the work environment and some user problems related to the new human-machine interfaces. Therefore, controllers verified constantly the data given by the system. Initially, the new control room was not an answer to the human demand. It was an adaptation of the controllers to the new system, during the first days of use, through particular field training method.

Those illustrative examples show concrete requests which can or may be encountered during the validation of a new working position organization, in a specific real context. A simulation are not free of critiques concerning the experimental protocol processing and entails some questions about the practical validation of the proposed system. Moreover, the field training case shows the choice problem of the means to be used to keep an optimal safety level, minimizing possible risks and combining economic and regularity criteria.

5. CONCLUSION

This article has discussed the study of human-machine organization and the case of air traffic control was detailed. A global methodology for designing and validating human-machine systems was proposed, using a working position representation. Working position communications were described for the command system organizations. Related to the air traffic control organization study, the paper has detailed the description of a control unit composed of working positions managed by human controllers for qualitative control. As a matter of fact, examples have illustrated the possible impacts of control tool changes in this work environment.

On the one hand, the adaptation of the human-machine system to the human operator is an important point during a system design. On the other hand, the human operator will have to adapt himself to the resulting system through specific and efficient methods for training on both simulated and real contexts. As a matter of fact, related to the present studies in air traffic control, controllers will work in a more and more virtual world which may modify the initial communication flow. Similar problems are already studied in human factors related to the cockpit in which pilots' training is more and more directed toward the crew resource management problems. Those studies results might be useful to orient those of the air traffic control about training methods or awareness programme to new human-machine system organization.

ACKNOWLEDGEMENTS

The author thanks the Milano-Linate CRAV (Italian acronym for Regional Center for Flight Assistance) for comments about Linate airport. He also thanks J. Vanderhaegen for his drawing contribution on the figure 6, M. Atkinson and S. Grant for their careful reading and corrections.

REFERENCES

Morico C. and S.L. Poirot-Delpeche (1992). Break or transition? A sociological viewpoint. *Control*, **70**, 18-19.

Salvi R. (1992). Process of design - The ten commandments. *Control*, **70**, 12-13.

Vanderhaegen, F. (1993). Multilevel human-machine cooperation between human operators and assistance tools - Application to the air traffic control. PHD thesis. University of Valenciennes, France.

Vanderhaegen F., I. Crevits, S. Debernard and P. Millot (1994). Human-machine cooperation: toward an activity regulation assistance for different air traffic control levels. *International Journal on Human-Computer Interaction*, **6**, 1, 65-104.

DESIGN AND EVALUATION OF AN ATC-DISPLAY IN MODERN GLASS COCKPIT

G. Huettig, A. Hotes, A. Tautz

Berlin University of Technology, Institute of Aeronautics and Astronautics
Section Flight Guidance and Control, 10587 Berlin, Germany

Abstract : To accommodate the future anticipated increase in air traffic, the Air Traffic Control (ATC) systems have to provide increased capacity by using enhanced technology. Future digital air-ground data link will enable the transmissions of automatically generated ATC messages into the cockpit. The feasibility of the visual presentation of such information is studied, using an Airbus A340 full flight training and research simulator. The study addresses in particular the development and integration of a display into the A340 cockpit, which presents tactical ATC messages and its evaluation. To specify the mental workload of the pilots, a secondary task technique, the NASA Task Load Index, a recall test, interviews and eye movement measurements are used.

Keywords : Air traffic control, Aircraft operations, Data transmission, Displays, Human factors

1. INTRODUCTION

The future anticipated increase in air traffic means that the Air Traffic Control (ATC) systems not only have to be harmonized, but have to provide increased capacity by using enhanced technology on ground and aboard in order to keep mental workload within manageable limits. One possible approach will be the implementation of air-ground data link, providing the Air Traffic Management (ATM) computers on ATC side with all flight information derived from aircraft on-board Flight Management Systems (FMS) (Bohr, 1990).

2. RESEARCH CAPABILITIES

The future ATM system and the integration of respective operational functions in the cockpit requires the optimization of the Human Machine Interface (HMI) in a highly automated aircraft environment. The A340 full flight training and research simulator at the Technical University of Berlin provides respective research capabilities. The Flight Management System Simulation (FMSS) provides a very close replica of the A340 FMS man machine interface. The system comprises the interactions with the Primary Flight Display / Flight Mode Announciator and Navigation Display, the Flight Control Unit, the Multipurpose Control and Display Units and could therefore be used to simulate the air/ground data link functions in a highly realistic scenario.

A very useful research tool of the simulator's Scientific Research Facility (SRF) is the Experimental Data Unit (EDU). The EDU allows the recording of an extensive number of different parameters during an inflight experiment. Not only the operational parameters, like the position of the aircraft, altitude or speed are recordable, additional parameters, e.g. concerning the use of the Flight

Control Unit (FCU), the occurrence of errors and the strength of sidestick movement can be recorded in real time during the simulated pseudo-realistic flight scenario.
Another important research capability is the ISCAN HEADHUNTER system, which allows the recording of the pilots field of vision at a rate of 50 Hz. Using this ISCAN system it is possible to monitor the direction of the pilot's field of vision, the precise point at which he looks, the pupil diameter and the fixation duration of different points in the cockpit.

3. METHOD

The project's objective is the development and evaluation of an experimental integrated ATC message display, where the visual presentation of ATC messages is integrated into the existing cockpit displays. This is realized using the development software TIGERS in combination with special RGB TERABIT monitors, which replace the original aircrafts navigational displays. The data exchange is implemented via a simulated data link between ground and aircraft. In particular, aspects of the human machine interface are addressed, e. g. acceptance procedures, visual perception issues and implications on the pilot's workload.

In a first step, the workload and scanning behavior including eye movement and fixation duration of pilots during flight in the A340 were studied. Therefore, a scenario with a flight from Munich (Germany) to Salzburg (Austria) and back was defined, using Standard Instrument Departure Routes (SID), cruise along the Air Traffic Service routes and Standard Arrival Routes (STAR). The pilots were asked to communicate conventionally via VHF with the appropriate ATC control units.

The objective of these tests was to provide initial data for the implementation of an ATC message display in a modern ´glass´ cockpit. The results of this first part of the study indicated, that the typical scanning cycle, which is defined by the number of viewing areas the pilot scans without scanning the same area again, includes only two different displays, before reaching a ´redundant area´.

This and other results, e. g. the fixation duration times, indicate, that it is not feasible to create an additional viewing area by implementing a new display for ATC messages (Huettig et al, in press). Therefore, an experimental integrated ATC message display was developed. The visual presentation of the ATC messages is integrated into the existing Navigational Display (ND), one of the Electronic Flight Instrument System (EFIS) displays of the A340, including the development of new acceptance and crew procedures.

In the main experiment, 10 experienced A340 airline crews participated. The pilots had to fly a standard scenario represented by a flight from Munich to Salzburg and back. Like in reality, the role of the crew as pilot flying (PF) and pilot non flying (PNF) was shared. Half of the experiments the pilots had to use the ATC display integrated into the Navigation Display for the transmission and presentation of the ATC message from ground to the aircraft. All messages from the pilots to the controllers were either transmitted via VHF and via pressing a confirmation button, respectively. For the manual confirmation of the displayed ATC messages, the rain repellent button was used, which is found on the lower end of the overhead panel approximately 35 cm away from the pilot's head.

As reference to today's flight situation, in the other half of the experiments the pilots had to fly the same scenario, but using only the conventional way of communication via VHF between the cockpit and the controller.

During the flights the autopilot was set inoperative for about 10 minutes on each leg. All other system's was normal during the experiment. The weather condition were above minimum with a low head wind of approximately 5-10 kt. The crew procedure and the acceptance procedure was practiced for approximately 5 minutes before the flights.

The implications of this data link environment on the pilots in flight compared to today's situation were investigated. In order to specify the differences between the two inflight situations during the main experiment, five different methods to measure workload differences were used.

The NASA-Task-Load Index (TLX) (Hart and Staveland, 1988) was used to evaluate the subjective workload experienced by the pilots. After each leg a TLX test was given to each pilot. During flight the pilots were ordered to estimate the time span of one minute. This secondary task technique was used to estimate the residual workload capacity. The pilots were asked to press the chrono-button in the glareshield in front of them every 60 seconds.

After the completion of the flight session, each pilot was asked to recall all of the ATC messages including their replies during the two legs flown. For further analysis the time used to complete an ATC message including the reply, aspects of the interview with each pilot and data of the ISCAN eye movement measuring system were used.

4. ATC MESSAGE DISPLAY

The experimental Navigation Display with an ATC Message Bar was developed using the existing Navigation Display with all possible modes and shrinking it down by 10 percent down. The ATC bar was displayed in the upper part of the ND. The experimental Navigation Display was rebuilt by using the graphic development tool TIGERS from the simulator manufacturer CAE of Canada. According to the judgments of the participating pilots, there was no noticeable difference to the original Navigation Display.

The following figure presents an example for a typical Navigation Display with an ATC message displayed.

Fig. 1. Experimental Navigational Display

When an ATC message occurred for the first time, a short chime was given and the text was displayed in yellow in the ND. If the pilot did not respond by pushing the ATC confirm button within a period of 10 seconds, the text started to flash with approximately 1 Hz. The ATC pictogram in the left corner of the ATC message display was presented for redundancy. After pushing the confirmation button, the ATC message stopped flashing and occurred in green color for about 10 seconds. After this period the ATC message was erased from the display.

5. CREW PROCEDURE

The crew procedure was developed to ensure complete knowledge, situation awareness and message information awareness for both pilots, the pilot non flying (PNF) and the pilot flying (PF). Distinction must be made between messages with immediate action required (change of ATC frequency, etc.), messages with delayed action required (e.g. request to report a future waypoint) and messages without action required (e.g. availability of a clearance). Every time when normally using the VHF for the communication between ground and aircraft, the PNF has to read back all the messages received from the ground. But if the message is transmitted via datalink to the aircraft and displayed in the ND, the PNF has to read the message to the PF. Then the PF has to confirm the message by the words „ATC confirm" or to refuse the message with „ATC refused". If the PF and the PNF want to confirm the message, the PNF has to push the „ATC confirm"-button. In all other cases the PNF has to use the VHF for the communication with the ATC controller and is not allowed to push the confirmation button, which in this stage of the study only has a status as a „read back and accept". If the PNF wants to inform ATC about the passing of a waypoint or has any other informations or questions, the use of the VHF communication is necessary.

6. RESULTS

The results of the subjective workload measurement TLX indicate, that flights in which the usual VHF communication was used, appeared to produce less workload than those flights in which the ATC message display was used. The TLX total scores are shown in the figure 2 below.

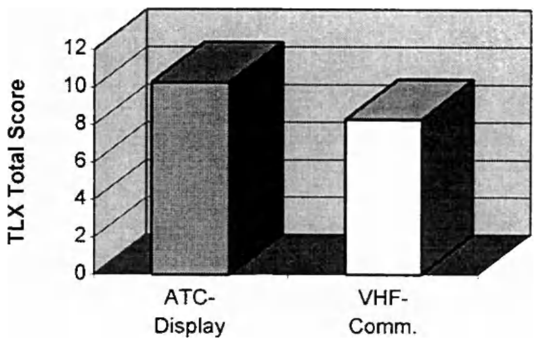


Fig . 2. TLX Total Scores

The differences between the VHF condition and the ATC display condition in total score were significant ($t = 2.29$, $df = 31$, $p < 0.05$). The analysis of the different factors of the TLX shows, that most of the differences in the total score can be explained by the two factors "Mental Demand" and "Temporal Demand".

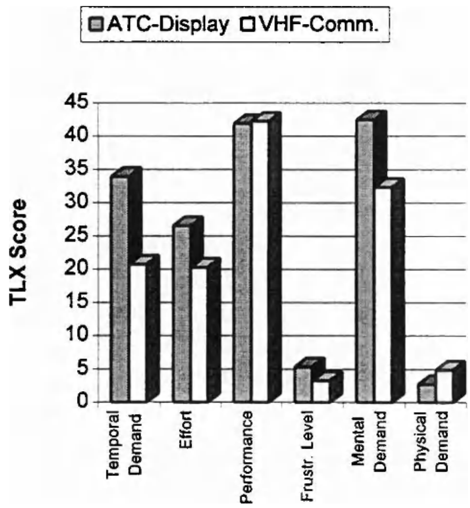


Fig. 3. TLX Factor Scores

For the factor "Temporal Demand" a significant difference can be found ($F = 4.37$, $df = 1$, $p < 0.05$). All the other factor differences are not significant.

Contrary to the subjective perception of the pilots, concerning the factor "Temporal Demand", the actual time spent to deal with an ATC message in the ATC display condition is less than the time used in the VHF condition. In average, the participating crews needed 2.75 seconds more to complete an ATC message, including its reply during the VHF condition. The average times of selected ATC messages are shown in the figure 4.

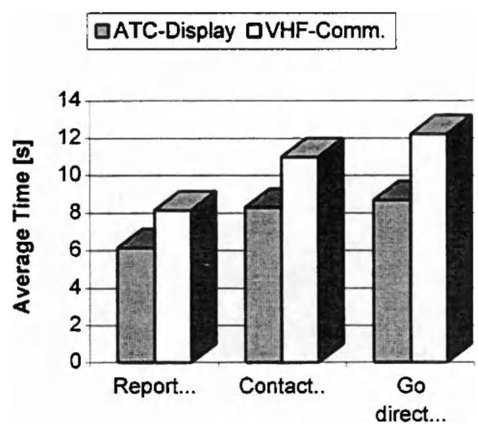


Fig. 4. Time to complete an ATC message

To analyze the amount of time used by the pilots to deal with each ATC message in both conditions, three typical ATC messages were selected. The message „REPORT ...", that e.g. orders the pilots to report a certain waypoint whenever the aircraft passes this waypoint, the message „CONTACT ...", which orders the pilots to contact a different ATC station and as a third message group the ATC message „GO DIRECT ..." was selected. This message orders the pilots to alter the flight path toward a given waypoint. For those three ATC message groups the time used to complete a message has been analyzed. The differences between the two conditions are not significant, except for the ATC message group of "REPORT..." ($F = 5.43$, $df = 1$, $p < 0.05$).

The results of the secondary task technique time estimation, shown in figure 5, are aggregated to an estimation index for secondary task measurements (Michon, 1966; Hart, 1975).

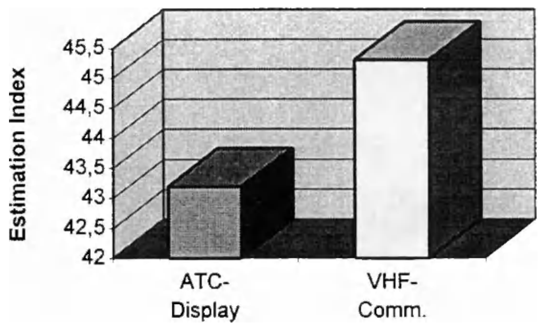


Fig. 5. Time Estimation Index

The scores indicate that in the VHF condition the time estimation was slightly more accurate, but there was no significant difference.

The results of the recall test after each complete flight session are shown in figure 6. The recall index was generated by rating each message with two points, if recalled completely, with one point if recalled partly and zero points if not recalled. The recall index that was measured in the ATC display condition is higher than in the VHF condition.

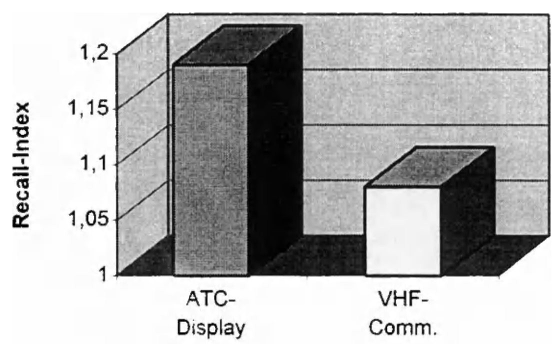


Fig. 6. Recall Index

The analysis of the eye movement measurements showed no differences concerning fixation time and cycle length between the two experiment groups and no differences compared to the first initial experiment (Huettig et al., in press).

The concerns of the participating pilots regarding the ATC message display were mostly related to the crew procedure and the amount of time used by this procedure. Since both pilots are reading the incoming message, it should not be necessary to repeat it by reading it out loud. The statement of the pilot flying to refuse or confirm a message should be enough to ensure the needed awareness. Besides, the pilots felt, that the new ATC display produced more workload, than the usual VHF communication. But considering the little overall time they used this message display, all of the pilots were convinced, that this implementation could be a useful way to display uplinked messages in the cockpit.

7. CONCLUSION

The feasibility of an ATC message display integrated into the cockpit environment of a highly automated aircraft - A340 - was demonstrated. Because of the future importance of air-ground data link in Air Traffic Management, some forms of presentation have to be developed, evaluated and also validated. The presented study is a first approach towards an integrated format of information presentation.

The results in regard to the workload measurements and to the eye movement measurements support this concept. Some improvements could also be identified in regard to time requirements for ATC communication.

However, it was also clear, that the optimization of the Human Machine Interface (HMI) and in particular the crew procedures are most important for future acceptance of such systems. Further studies will therefore concentrate on HMI aspects and a further integration of the aircraft Flight Management System to e.g. direct alterations of flight plans and trajectories in the process of the communication through data link.

REFERENCES

Bohr, T. (1990). ATM-Kooperativer Weg in die Zukunft. *Symposium Auswirkungen neuer Technologien auf die Sicherheit im Luft-verkehr.* TÜV Rheinland.

Hart, S. G. (1975). Time Estimation as a Secondary Task to Measure Workload, In: *Proceedings of the 11th Annual Conference on Manual Control*, (Report No. NASA TMX-62). Washington, DC: U.S. Government Printing Office, pp. 64 - 77.

Hart, S.G. and Staveland, L.E. (1988). Development of NASA - TLX (Task Load Index) : Results of Empirical and Theoretical Research. In: *Human Mental Workload* (Hancock, P.A. and Meshkati, N.). North-Holland, Amsterdam.

Huettig, G., Hotes, A. and Tautz, A. (in press). Human Factor Design Considerations for Air Traffic Control Information Displays in a Modern Glass Cockpit. *Proceedings of the 21 st conference of Aviation Psychology*, EAAP, Dublin.

Michon, J. A. (1966). Tapping Regularity as a Measure of Perceptual Motor Load. *Ergonomics*, 9, 67-73.

ANALYSIS AND MODELING OF FLIGHT CREW PERFORMANCE IN AUTOMATED AIR TRAFFIC MANAGEMENT SYSTEMS*

Kevin. M. Corker[1]
Gregory M. Pisanich[2]

NASA Ames Research Center[1]
Moffett Field, California

Sterling Software Inc.[2]
Palo Alto, California
USA

Abstract: NASA has recently initiated a research program to enhance commercial aviation throughput in crowded terminal areas. The Terminal Area Productivity (TAP) initiative seeks to focus technologies in automated air traffic management and to improve flight management systems. The interaction between air traffic control (ATC) automation, flight deck automation and the human operators who mediate them will be reported here. We've performed empirical studies in full-mission simulation to address procedural coordination between flight crews and air traffic controllers. In parallel, we have addressed the need for rapidly adaptable analytic techniques by developing a human performance model of flight crew.

Keywords: Air Traffic Control, Aircraft Operations, Modelling, Computer Simulation.

1. INTRODUCTION

The operational environment in the evolving national airspace is changing the nature of the tasks of flight management and airspace management. Task analytic data reveal that the flight crew (increasingly a flight-management crew) spends the majority of its time in cockpit supervising the automated systems that are flying the aircraft and communicating with the air-traffic management systems that are directing the aircraft through an increasingly crowded airspace (McGuire, Zich, Goins, Erikson, Dwyer, Cody and Rouse, 1990; Wiener and Nagel, 1988). As automation advances in ATC, the same can be said of the air traffic controller. Flight control technologies have provided automatic control that is more timely, more fuel efficient, and, if properly implemented, more reliable than human pilots in normal operation. Air traffic automation is providing support and decision aiding for the controller that will guide the controller to performance that would be unmanageable without automated aiding (Erzberger and Nedell, 1989, Davis, Erzberger and Green, 1991, Erzberger, Davis and Green, 1993). In terms of airspace management philosophies, the goal of increased *terminal area* productivity is expected to be obtained through tighter and more continuous control of the terminal area airspace (Williams and Green, 1993 and den Braven, W. 1992). Parallel independent simultaneous runway use under instrument meteorological conditions (IMC) , curved approach trajectories, and decreased trailing and lateral distances all suggest a system wherein the margin for error is very small and the propagated effects of error or anomaly are potentially very large. We have applied empirical tests and models of human performance to begin to examine the procedural impact of the introduction of automation aiding for ATC on the operation of the flight deck.

2. CENTER TRACON AUTOMATION SYSTEM (CTAS)

The clearances that are provided to the flight crew in our full mission simulation are generated by the Center TRACON Automation System (CTAS) developed at NASA Ames Research Center in collaboration with the FAA. CTAS consists of three types of automated advisories for controllers in management of the flow of traffic into the terminal area. The traffic management advisor (TMA) generates landing sequences, runway assignments and landing times for all arriving traffic. The descent advisor (DA) generates clearances for controllers guiding aircraft from the enroute portion of their flight into the feeder gates. The descent profile provided by the DA includes conflict free descent and arrival/crossing times in descent. The DA can also specify a top-of-descent (TOD) point to insure fuel-efficient flow of traffic into the final approach. (The operation of the DA is the portion of CTAS that we have concentrated on in this study.) The Final Approach Spacing Tool (FAST) provides speed and heading advisories for amendment to approach control. The integrated operation of three tools provides a critical technology to realizing the benefits of the TAP program.

2.1 DA and Operational Requirements: Flight Deck

The descent advisor calculates in fast-time, the equations of motion for an aircraft approaching the terminal area. It provides advisories for cruise speed, TOD, and descent speed profiles. The flight crew attempts to meet the requirements of this descent profile through the flight management technology available in the specific aircraft being flown. (The aircraft simulator flown in our study were equipped with current generation flight management systems.) Timeliness and accuracy in compliance with the CTAS advisories is essential to maintaining a safe/efficient profile in descent.

The flight crew faces several constraints in complying with and enacting the CTAS clearance. They must decide if there is time to enact the clearance from the point of receipt to the point of TOD. They must consider the flight characteristics of their aircraft. They must consider the meteorological situation, the comfort of the passengers, and any anomalous conditions that might constrain their ability to comply. Examples of weather factors would include reported icing, turbulence, or visible connective weather. Equipment factors include any equipment problems or descent limits. Factors influencing passenger comfort include pressurization limits, high turbulence, or an excessive rate of descent. In addition they must decide how (using what flight management tools) they intend to comply with the clearance, and then enact the proper interaction with their equipment to assure successful compliance.

That process of decision and action has been the focus of our analyses and modeling. In making this decision, these rules are instantiated with information gathered from ATIS (Automated Terminal Information Service) and ACARS (Aircraft Communications Reporting and Addressing System) or from observed weather or pilot reports from preceding aircraft. The information gathered is applied to decision criteria that are applied as rules or thumb or fuzzy rule sets, not as hard limits. Some of this information is used to program (or inform) the Flight Management System (FMS), which generates a fuel-optimized top of descent point, descent profile, and speed for that aircraft .

2.2 Air Traffic Control

As the aircraft approaches, CTAS also performs its own calculations attempting to integrate the subject aircraft into the overall flow of traffic. The more information about current traffic conditions, crossing traffic, as well as the subjects aircraft state that CTAS has the better its prediction. CTAS as an optimizing system would prefer to issue a TOD clearance to a subject aircraft fairly late in that aircraft's approach to the TOD thereby providing the system more time and optimization cycles. and incurring less variability in a flight crew's response. The CTAS clearance is generally issued as an amendment to the profile descent that the flight crew has previously entered into the FMS. This clearance also specifies a top of descent point (distance from some way point), descent speed, bottom of descent altitude, and bottom of descent speed, all of which may be different from those generated earlier by the FMS. This difference derives, in part, as a result of incomplete information in the CTAS system relative to immediate aircraft state variables.

2.3 Compliance

Given the time available, the flight crew must make an initial determination as to whether a reasonable decision on the received clearance can be made at all. The flight crew takes an initial look at the clearance and makes quick, time/distance judgment based on prior experience as to whether the aircraft can comply with the clearance. After communicating their intent to accept, the flight crew then selects the appropriate automation method for implementing the clearance. This decision is based on an estimate of the time remaining to the top of descent and the implementation preferences of the flight crew. Although encouraged to use the highest level of automation possible (Vertical Navigation or VNAV), when time is short (because of delays or excessive distractions), the flight crew will forgo entering the clearance into the FMS and begin the descent using a more manual mode (for instance, Flight level Change or FLCH). Once the descent has begun, the information can be entered into the FMS and the higher mode activated.

3. HUMAN PERFORMANCE MODEL

We have adapted the Man Machine Design and Analysis System's representation of a human operator to this top of descent clearance scenario (Corker and Smith, 1993). This framework provides a collection of models that describe (within the limits of the accuracy of the constituent models) the responses that can be expected of human operators in several areas that are critical to safe and reliable operation of advanced aircraft systems. To provide a relatively complete and useful representation of the human operator in these systems, we need to account for three aspects of the operator's behavior: perceptual processes, cognitive processes, and response processes. Each of the human operators modeled by MIDAS contains models and knowledge-base structures, the interaction of which will produce a stream of simulated operator activities in response to mission requirements, equipment requirements, and models of human performance capabilities and limits. Figure 1 illustrates the components of the full MIDAS model. Due to the modular nature of the MIDAS framework not all of the modules depicted need be active in any particular application. The model elements highlighted in the figure are the models that were used in the commercial aviation scenario and they are described in some further detail below.

Air MIDAS

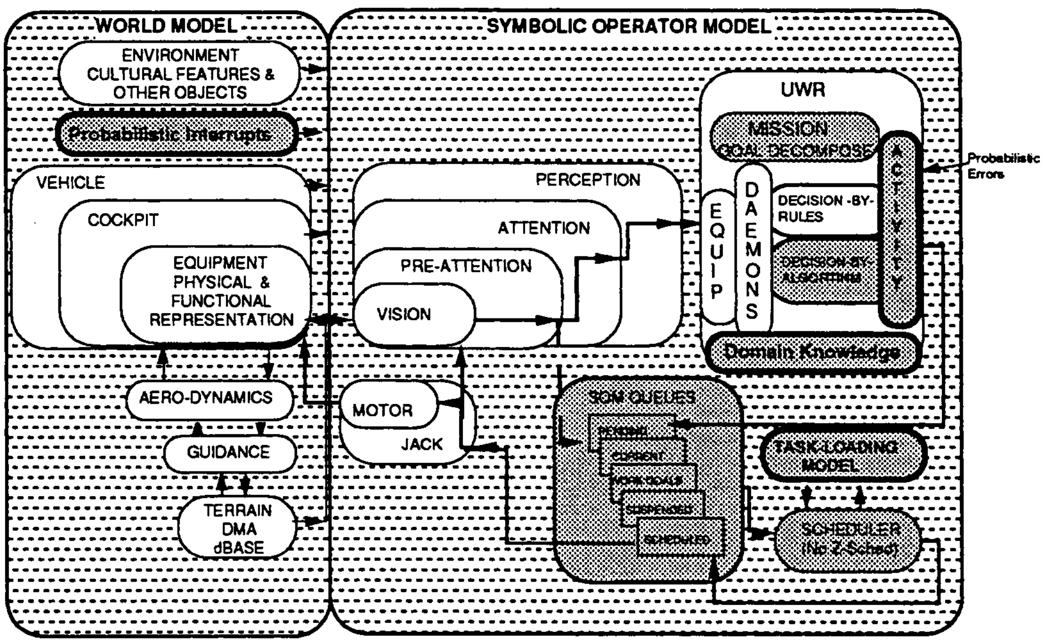


Figure 1. Full MIDAS closed loop model with all modules represented. Modules active in this simulation are shaded.

Updatable World Representation (UWR) In MIDAS, the Symbolic Operator Model, or SOM, provides a mechanism whereby human agents representing individual and potentially cooperating teams of pilots and controllers accesses their own tailored or personalized information about the world. This internal representation of world knowledge is called an UWR (Updatable World Representation). The contents of an UWR are determined, first, by pre-simulation loading of required mission, procedural, and equipment information. Then data is updated into each operator's UWR as a function of the perceptual mechanisms modeled in the operator. The data of each operator's UWR is operated on by daemons and rules to guide behavior and are the sole basis for a given operator's activity. The individual operator may, or may not, receive a piece of information available to the sensory apparatus as a function of perceptual focus and activity load at the relevant point in the mission. It is of some significance that, while ideally the human operators' representation of the world would be consonant with the state of the world, in fact this is rarely the case. The capability for both systematic and random deviation from the ground truth of the simulation world is a critically necessary component of any system that intends to represent and analyze nontrivial human performance. In this simulation the activity of the crew is guided by application of a set of rules to the receipt, acceptance/rejection, and enactment of the ATC CTAS clearance.

Activity Representation: Activities are MIDAS objects that simulate actions performable by an agent in the system. Representations of activities available to an operator are contained in that operator's UWR. Activities are organized by flight-system and mission goals. Activities in this scenario are characterized by: *preconditions* that define the allowable conditions for their initiation *satisfaction conditions* which define their successful completion; *spawning specifications* which detail the temporal and logical constraints on any "child activities" that might be needed for activity performance; *decomposition methods* that describe in a context-sensitive way what children should be spawned to accomplish a higher-level activity's goals; *interruption status and interruption specifications* which detail the interruption and resumption methods for that activity; and *duration* either estimated or calculated by an activity-specific function. These activities are the drivers for simulation action and are recorded as a function of simulation runs. The rules that the simulated air crew apply to a decision for action at the TOD are actions in the sense described above wit the added element of contingency represented in a propositional structure. The data to decide the proposition are found in the updateable world representation of the flight crew member or are sought by perceptual interrogation of the external simulation world

4. EMPIRICAL STUDY AND MODEL ANALYSES

An empirical study was performed in a full mission simulation at the NASA Ames Crew Vehicle Simulation Research Facility in order to examine the procedural impact of the interaction of CTAS-base clearances on flight crew operations in interaction with their flight management systems.

4.1 Human-in the-Loop Simulation Method:

The study was performed in a certified flight simulator of a 747-400 Boeing aircraft. The simulator is equipped with a current generation flight management system and all active systems and subsystems of that generation cockpit. The simulator is on a moving base that recreates flight action in six axes of motion. In addition to the aircraft systems out-the-window views are supplied that include terrain, airport and other aircraft in the airspace. The subjects in this study served as volunteers and were all current line pilots from a single airline certified in the 747-400. Four two-pilot (Captain and First Officer) crews participated in the study and each of the pilots made two approaches and one landing in the experiment. The result was eight runs per flight crew across the experimental manipulations and sixteen runs for the human pilots overall. The experiment paradigm was a within subjects design in which each of the crew members experienced varied conditions of flight in their two descents. In order to investigate the impact of variations in CTAS clearances on FMS operation, the experiment varied three factors. First, starting from a cruise speed of 300 knots indicated airspeed (KIAS) the required speed in descent was given at two levels (260 or 320 KIAS). Second the altitude at which the descent was to be terminated was varied among four altitudes (17, 19, 21 & 23 thousand feet). Speed at this "crossing restriction" was required to be 250 knots. Finally, in order to press the crew relative to amount of time they had to enact the clearance, the distance before the TOD at which the clearance was provide was varied between 15 or 20 nautical miles prior to TOD. The flight crews were "inserted" into the airspace over Denver Stapleton International Airport at 33,000 ft. and asked to fly one approach to the crossing restriction ("DRAKO") and one down to land at a simulation of Denver.

4.2 Model-based Simulation Method:

In order to compare the performance of the predictive model of flight crew behavior with that of the human flight crew, the model was initialized as the human flight crew had been. The model flight crew was then directed to fly descent profiles with the same conditions of speed, crossing restriction, and distance to top of descent as experienced by the human flight crews. We used a split-halves method in which the model's activity times (e.g. FMS operation and "button-push" time) were derived from one half of the human performance data and the model was tested to see if it could predict the behavior of the remaining half of the human performance data. The model has a stochastic element of interruption in communication to replicate communication with ATC. The interruption data were also generated by a split halves method so that the frequency and length of interruption that the model is providing is based in a reasonable set of actual ATC operations. The model invokes a contingent decision activity in which it considers, through its rules, whether or not a clearance is acceptable and when and how to enact it using the variety of methods within its activity set. The model was run eight times to match the human performance runs and its stochastic property was tested by performing four sets of eight runs. The model generated activity data comparable to that generated by the human operators. For example, data collected for the human flight crew included time to complete access and to enter the clearance, time to receive the clearance, time to decide on clearance acceptability, ATC communication times, and FMS operation times. The model similarly generated activity times for that sequence of behaviors that can be used in direct comparison. It should be noted that at the present time the model represents the activity of one crew member. In this case that activity of the pilot-not-flying. Flight crew cooperative activity is not presently represented, though that capability is under investigation..

5. RESULTS

The experiment and model are being used to analyze the procedural impact of ground and aircraft based automation. That analysis will include examination of data-link and voice operations and a detailed analysis of the effects automation interaction on the procedural sequence. The purpose of this paper is to provide evidence of the efficacy of the overall approach of a model-based analysis of complex human performance in the flight deck environment. With that purpose in mind, we confine our analysis here to a comparison of the model behavior to that of the human-in-the loop simulation, a verification of overall model performance. (The impact of the specific CTAS manipulations on flight crew performance will be reported in an upcoming NASA Technical Memorandum.) The hypothesis being tested then is that the model performance across the manipulations of CTAS will not be significantly different from that of the flight crew in the same performance regimen. There are multiple measures available to test flight crew versus model performance. We will compare model performance in the critical flight phase of receiving the CTAS TOD clearance, deciding whether and how to enact it and preparing the aircraft systems for the required TOD. This performance measure is being used in our ongoing analysis of automation integration issues. The variable reported here is the time remaining between the aircraft being fully configured for descent and the TOD point. This is a measure of the "spare time" the crew has between receiving and enacting a clearance and its required time of completion, i. e., the TOD point. This measure is applied across the conditions of the experiment for comparison of model and human behavior.

Three independent comparison were made and t tests were applied. First a comparison was made between the behavior of the model across the experimental conditions and the performance of the flight crew across those conditions (split-halves comparison). Second a comparison was made between the one model run and cumulative data from the four simulation model runs to check for internal consistency in the model (cumulative model data versus single run). Finally, the accumulated human data was compared to a single model run, chosen at random, to see if there was an effect on the model variance encountered by summing across model runs (flight crews versus simulation run.) In all cases the t test (df = 14) revealed no significant differences between the data sets compared. The comparisons are presented in Table 1.

Table 1. Three model versus human performance data comparisons

	t	p (T<t) two-tailed
split-halves	1.286	0.23
cumm. model /single run	1.088	0.29
flght crew /single run	1.017	0.33

6. DISCUSSION

The data suggest that the model performance is predictive of the flight crew performance across the conditions of this experiment. The lack of significant difference between the model and the cumulative human performance and the check for internal consistency in this model are encouraging to our effort. The simulation model produced a stream of activity that included decision and interaction with varied levels of flight management automation. In the time-stressed operation at TOD, the model behavior is predictive of the flight crew's behavior in terms of the time taken to meet an ATC profile descent clearance.

The particular effect of the experimental manipulations on the flight crew procedures and on the model's replication are not tested in this cross comparison. The effect of time-stress on performance is likely to be seen in procedural sequence changes and inter-activity time shifts which we have not include in the present analysis. The model supports such a comparison and given its initial positive indication of as a predictor of overall crew performance we will begin to examine its effectiveness in representing the micro-structure in the "strategic differences" of enacting a particular clearance. It will be our next task to examine the impact of a change in media of presentation of the clearance (voice versus data-link). It is also of interest to examine the potential procedural changes required by a misalignment of the ground -side and flight deck automation and the crew's requirement to bring them into concurrence. The selection of flight modes and their efficiency are also being examined.

The model of human behavior in its coarse level predictability provides a tool for efficient and effective examination of these issues. Verification of the model's micro-behavior accuracy is required before the model-based analysis can be considered as a fully credible design-analysis tool. We are pursuing that verification.

REFERENCES

Corker, K.M., and Smith, B. (1993). An Architecture and Model for Cognitive Engineering Simulation Analysis: Application to Advanced Aviation Analysis. Presented at AIAA Conference on Computing in Aerospace. American Institute of Aeronautics and Astronautics, San Diego, CA.

Davis, T.J., Erzberger, H. and Green, S. M. (1991). "Design and Evaluation of and Air Traffic Control Final Approach Spacing Tool." Journal of Guidance Control and Dynamics Vol. 14 (4). July-August.

den Braven, W. (1992). Design and Evaluation of an Advanced Air-Ground Data-Link System for Air Traffic Control. NASA Technical Memorandum (TM-103899) January.

Erzberger, H. and Nedell, W. (1989). Design of an Automated System for Management of Arrival

Traffic. NASA Technical Memorandum (TM-10220) June.

Erzberger, H., Davis, T. J. and Green S.M. (1993). Design of Center-TRACON automation System. AGARD Guidance and Control Symposium. May: Berlin, Germany.

McGuire, J. C., Zich, J. A., Goins, R. T., Erikson, J. B., Dwyer, J. P. Cody, W.P. and Rouse, W. (1991). Functional Decomposition of the Commercial Flight Domain for Functional Allocation. NASA Contractor Report NAS1-18028.

Williams, D.H and Green, S.M. (1993). Piloted Simulation of and Air-Ground Profile Negotiation Process in a Time-Based Air Traffic Control Environment. NASA Technical Memorandum (TM-107748) April.

*ACKNOWLEDGMENT: This work was supported by NASA Ames Research Center and the U.S. Army Aeroflightdynamics Directorate: Army/NASA Aircrew Aircraft Integration Program (A^3I).

ENHANCED VISUAL DISPLAYS FOR AIR TRAFFIC CONTROL COLLISION PREDICTION

MEGAN JASEK*, NICHOLAS PIOCH * and DAVID ZELTZER*

*Sensory Communication Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

Abstract. Conventional air traffic control (ATC) displays are simple 2D overhead views using symbols and text to represent aircraft and their altitudes. To improve accuracy and timeliness of predicting aircraft collisions, several alternate 2D and 3D ATC displays involving multiple views and stereo perspective graphics were designed. Results from a series of experiments involving scenarios with varying numbers of aircraft indicate that a 2D graphical display with multiple views presents the information more effectively than the 3D displays or the conventional 2D display.

Key Words. Air traffic control; Computer graphics; Displays; Stereo vision; View angles

1. INTRODUCTION

Current air traffic control (ATC) displays are 2D overhead views that show the relative ground position of each aircraft graphically, while the altitude is displayed numerically. Controllers must predict potential collisions between aircraft by monitoring their X-Y positions and their changes in altitude. Deciding whether the aircraft will collide in the X-Y coordinate plane is straightforward with the visual cues of the moving graphical icons, but determining if the aircraft will be at the same altitude when their paths cross is a much more difficult task with only the numerical information displayed.

The motivating hypothesis of this work was that providing more visual feedback for the relative altitudes of the aircraft would enable the controller to make faster, more accurate decisions about whether or not aircraft will collide. Initial experiments, however, confirmed results reported by Ellis and others, namely, that increasing an operator's "immersion" or sense of "presence" in a given visual task does not necessarily enhance performance, but may in fact hinder it (McGreevy and Ellis, 1986; Barfield and Kim, 1991). In the collision prediction task studied here, one can argue that a properly designed 2D display of 2D data presents the operator with a concise view of only the essential data required to make decisions; such a 2D display has abstracted out a wealth of information that is irrelevant to the task. Given an interactive perspective view of a 3D scene, however, perhaps viewed in stereo, an operator first has to find an appropriate viewpoint and direction of view so that altitude, distance and course estimates can be made. Not surprisingly, results here show that this added view control task does indeed degrade performance.

In these first experiments subjects were asked to make collision predictions regarding pairs of aircraft. Later experiments probed how increasing the complexity of the task by increasing the number of aircraft in the scene would affect operator performance with the 2D and 3D views. In order to factor out the view control task in the 3D cases, two new displays that used automatically-computed, non-interactive viewpoints were introduced.

The purpose of this work is not solely to develop better displays for predicting straight-path mid-air collisions. After all, computers can certainly perform this type of mid-air collision calculation more quickly and reliably than humans (Harman, 1989). However, the results reported here may be relevant to human ATC collision-prediction performance under more realistic scenarios in which aircraft traverse curved and dynamically changing paths. Moreover, this work represents an exploration of the more general problem of data visualization.

Section 2 describes the initial experiment involving collision prediction between pairs of aircraft. Section 3 discusses the results of this first experiment. Section 4 describes a similar experiment involving three aircraft. Section 5 describes the results of Experiment 2. Section 6 describes informal results from further experiments involving more than three aircraft. Section 7 offers concluding remarks on the overall research.

2. DESIGN OF EXPERIMENT 1

For the first series of experiments, four different visual displays were designed, each showing ten scenarios of two aircraft flying in straight line paths with constant velocities. All experiments used a continuous update rate with zero latency, rather than the four-second radar scan time used in actual TRACON facilities. For each scenario the subject must predict whether or not the two aircraft will collide. The conditions signifying a collision are that the altitude difference is less than 50 feet and the differences in X and Y coordinates are both less than 100 feet. For each decision that the subject makes, the following data are recorded:

1. whether the subject predicted a collision or a miss, and
2. how many seconds before the collision (or pass-by) the subject made the prediction. In the case of a miss, the time before the instant of closest approach is recorded.

The subject must press the space bar if he or she believes the two aircraft will collide and the control key if they will miss. The subject is told that for each scenario a score is given, calculated by assigning a point for each second before the collision or pass-by that the subject made the prediction. Incorrect decisions receive a score of zero (although the timing of the decision is still recorded, for a more complete analysis). Thus, it is advantageous to predict as early as possible, but only if the subject is reasonably sure of the correctness of the prediction. In each display the subject receives feedback both to acknowledge the pressing of a prediction key and to notify that a collision has occurred.

Training scenarios were provided at the beginning of each display version to help negate the learning effect. Moreover, the order in which the four versions were presented was randomized for each subject, to counteract learning trends across versions. In order to compare individual scenario scores across different versions, the same ten scenarios were used for each of the four versions. Subjects were never aware of this, since the differences in appearance of each of the displays were dramatic and the ordering of the ten scenarios was randomized for each display. After completing the scenarios for all four versions, comments were recorded about which displays were most and least helpful to the task.

2.1. *2D Version (2D1)*

The 2D version is a simplified model of a current ATC radar screen that shows an X-Y plot of the aircraft along with their current altitudes. This version consists of a grid of green lines on a black background, denoting intervals of 1000 feet. Aircraft are displayed as triangles moving across the screen, pointing in the direction of flight. Each triangle has a box moving alongside it, displaying the altitude of the aircraft (see Figure 1). Since the area displayed was rather large (one inch corresponded to approximately 1000 feet), the icons for two aircraft meeting the horizontal collision criterion would necessarily be touching, thus proximity in the X-Y coordinate plane is easy to detect. However, due to the rapidly changing numerical fields, trying to predict whether the vertical collision criterion would occur is significantly more difficult.

2.2. *2D Split-Screen Version (2DS)*

This version splits the screen into two parts. The left side of the screen is the same X-Y plot as the 2D1 version above. The right side shows an X-Z vertical cross section of the same scene, so that the subject can graphically see the relative altitudes of the two aircraft. Altitude numbers are displayed on the left side only (Figure 1). To further help the subjects match up icons on the left side with their counterparts on the right, each aircraft is displayed with a different color. Unlike the 2D1 display, the right side allows subjects to observe coincidences in height by checking whether the aircraft are graphed at the same Z coordinate. Thus, they could use the following simple rule of thumb to predict collisions: if the two icons are headed toward each other on both screens, the aircraft will collide; if the icons will miss each other on either one of the screens then the aircraft will miss.

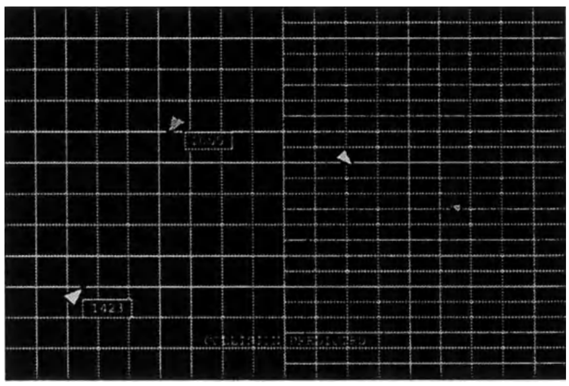


Fig. 1. Display version 2DS. Version 2D1 displays only the left side.

2.3. *3D Version (3DV)*

This version displays the aircraft in a 3D virtual environment, rendered on a Silicon Graphics Onyx workstation. The subject wears field sequential stereo glasses (StereoGraphics' CrystalEyes) to see the image in three dimensions. A red aircraft and a blue aircraft are shown flying over a flat ground (Figure 2). The subject can change the view angles by dragging with the left mouse button vertically or horizontally. Dragging with the middle mouse button provides a zoom feature, so that the subject can move in closer to the scene. In the 3D versions the altitude is not displayed numerically, ensuring that subjects rely only on the visual cues unique to a three-dimensional display. The appearance of collisions in this display is more natural in that if any part of one aircraft touches any part of another, a collision has occurred.

2.4. *3D Version with Drop-Bars (3DB)*

This version is the same as the previous one except that a perpendicular drop-bar was added to each aircraft to provide additional depth cues for altitude estimation (Durfee and R. M. Willis, 1992). The drop-bar extends from the bottom of the aircraft to the ground and is segmented with different colors every 200 feet. This feature is intended to help the subject perceive the relative altitudes of the two aircraft as well as their ground positions (Figure 2). The viewpoint is controlled in the same manner as in the 3DV version above.

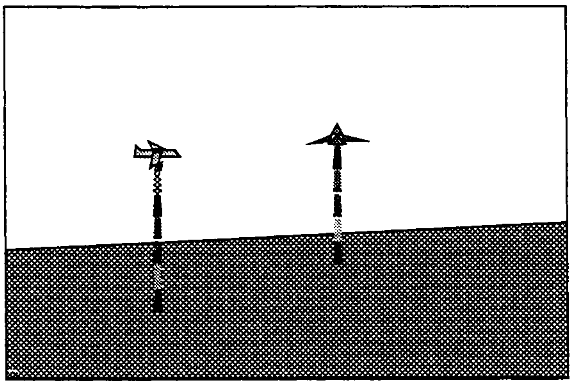


Fig. 2. Display version 3DB. Version 3DV is the same, but without drop-bars.

3. RESULTS OF EXPERIMENT 1

Thirteen subjects participated in Experiment 1. Their average scores are represented by the solid line in Figure 3. Contrary to what was expected, the average scores for the 2D versions were almost 2 points above the 3D versions, with the 2DS version at the top. In some scenarios when a miss was obvious in the X-Y or X-Z projection planes, a subject could quickly make a correct decision in the 2D versions because these projection planes are isolated in these displays, explaining why some scenarios score much higher in the 2D versions than in the 3D versions (Figure 4, scenarios 2, 4 and 5). On the other hand, for scenario 3, the subjects scored consistently higher in the 3D versions than in the 2D versions (Figure 4), perhaps because subjects found that viewing the scene from multiple angles was helpful.

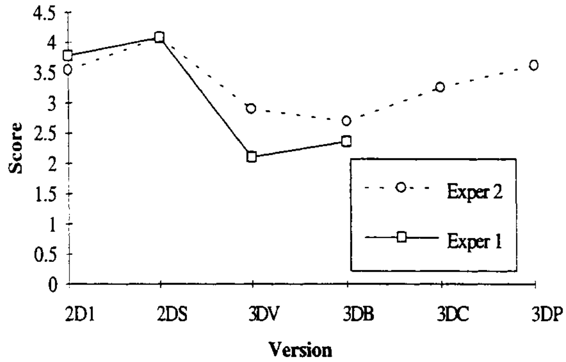


Fig. 3. Overall average score by version for Experiments 1 and 2.

Note that the score provides a value that represents both timing and accuracy for the subject. Accuracy alone was about the same for all four versions with approximately 20% incorrect in each. Given this uniformity in prediction accuracy across all versions, the results indicate that subjects were able to arrive at more timely decisions with the 2D displays. The cause of this slower performance in 3D is most likely attributed to struggles with the additional interface task of finding a good viewpoint with the mouse. A related issue concerns the strategy used by the subjects to make collision predictions. Many subjects used projections of aircraft onto coordinate planes to judge altitude, course, and proximity, as evidenced by a frequent alternation between vertical and horizontal view angles in the 3D versions. Those who did not exhibit this switching of views in the 3D versions would have to perform these coordinate plane projections in their heads. Either way, these subjects would do better with the 2D displays, in which this information is presented directly, than with the 3D displays, leading to increased times for decision making. Whatever strategies were used, however, the results show that the time is used wisely for the 3D versions since their accuracy is similar to that of the 2D versions.

To determine if the 2D and 3D scores were significantly different, a series of t tests were performed. A matched-pair t test was used to compare each 2D version to each 3D version at a significance level of 95%. The tests showed that both the 2D1 and the 2DS versions had significantly higher mean scores than the 3DV and 3DB versions.

The subjects' comments consistently identified the 2DS version as the easiest and the most effec-

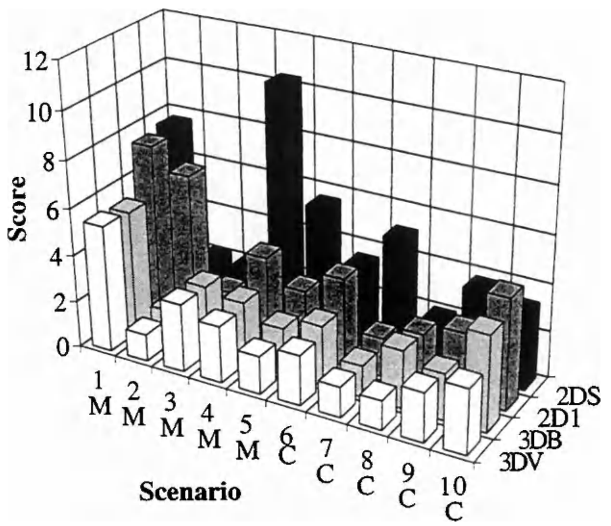


Fig. 4. Average scores for each scenario by version in Experiment 1. "M" denotes a miss; "C" denotes a collision.

tive display. Subjects often remarked that using the mouse to change the viewpoint in the 3D displays was difficult, but could possibly become easier with more practice. There were mixed reviews about the drop-bars in the 3DB version; approximately half of the subjects liked them and half did not.

4. DESIGN OF EXPERIMENT 2

To determine the value of the above displays for more complex situations, a new set of nine scenarios using three aircraft was designed. Two new 3D displays were implemented for inclusion in the new experiment in addition to the original four display versions. These new displays featured computer-controlled viewpoints and view angles in order to eliminate the view control interface task. However, adding a third aircraft to the scenarios led to four possible collision outcomes instead of just one. To avoid the task complexity of specifying exactly which aircraft will collide and to allow more valid comparisons of new results with the results from Experiment 1, the binary choice method of Experiment 1 was retained. Thus, the subject would press the space bar if any or all of the three aircraft were thought to collide and press the control key if all were thought to miss.

4.1. *Automatic Viewpoint Algorithms*

Transferring viewpoint control from the subject to the computer required an automatic computation of an optimal viewpoint for collision detection based on the positions of the three aircraft. Several attempts at this had been made earlier for Experiment 1 using two aircraft. One idea was based on Das' thesis on finding optimal viewpoints for teleoperator displays. Using Das' technique here, the computer placed the viewpoint on the geometric plane orthogonal to the line between the two aircraft and intersecting the midpoint of that line (Das, 1989). The continuous nonlinear movements of this algorithm proved to be quite confusing, making it hard to isolate a direction of flight for the two aircraft. To alleviate this confusion, the same algorithm was tried again with discrete updates in viewpoint at four-second intervals. This version was more suitable for isolating directions of flight, but caused disorientation after each periodic "cut." Based on these failed attempts, it was decided that the best algorithms would involve a simple linear motion of the viewpoint at a constant velocity, thereby minimizing any confusion between viewpoint motion and aircraft motion.

4.2. *3D Cockpit-Centered Version (3DC)*

This new display places the viewpoint in the center of one of the three aircraft (called the "owncraft"). The view angle is updated dynamically to orient along a ray that bisects the angle between the two rays from the owncraft to each of the other two aircraft (Figure 5). This direction is most likely to include both of the other aircraft in the field of view. To further ensure that both of the other aircraft will be visible, the field of view is increased from the usual 45 degrees to 90 degrees. Additional objects are drawn on the ground, such as a runway and simple buildings, to help the subjects discern the owncraft's direction of flight relative to the view direction. Also, the owncraft is not drawn since it could obscure large portions of the field of view. Before beginning training on this display, the subject is reminded to watch both for collisions between the owncraft and another aircraft, as well as collisions between the other two aircraft. For the former case, the "looming" cue, in which an aircraft remains in the same location on the screen but grows larger with time (Levison *et al.*, 1994), is suggested to the subject. Finally, the subject is reminded that the view direction is not necessarily the sâme as the direction of flight and may be constantly changing to stay focused on the other aircraft.

4.3. *3D Programmed External View (3DP)*

This display provides a computer-controlled external view of all three aircraft, zooming in continuously as they approach each other so that they span the entire field of view at all times. The algorithm for computing the viewpoint is as follows. One aircraft (usually the lowest in altitude) is chosen as the center of focus. At all times the

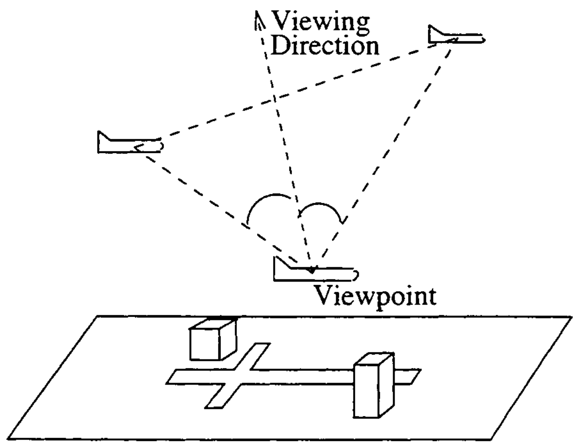


Fig. 5. Viewing algorithm for 3DC version.

viewing direction will be oriented toward this aircraft, thus it will always appear in the center of the screen. The viewpoint is chosen by finding the ray extending from the focus aircraft and bisecting the line segment connecting the other two aircraft. The viewpoint is placed on this ray at two and a half times the distance from the focus aircraft to the connecting line segment (Figure 6). Thus the viewpoint will always lie in the unique geometric plane defined by the three aircraft, and the view angle always centers on one aircraft, keeping the other two on opposite sides of the screen. Since altitude differences at the start of each scenario are always small compared to horizontal distances, this algorithm begins from a relatively horizontal vantage point, which makes differences in altitude readily apparent. As the scenario progresses, the viewing angle often migrates toward an oblique angle from which the horizontal motion can also be readily discerned.

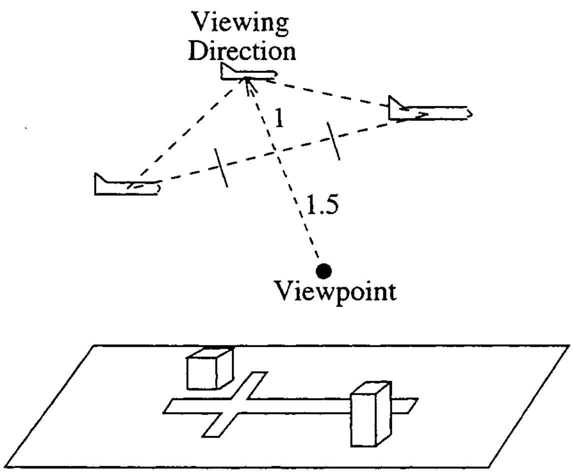


Fig. 6. Viewing algorithm for 3DP version.

5. RESULTS OF EXPERIMENT 2

Eleven subjects participated in Experiment 2. Their average scores are represented by the dotted line in Figure 3. Again the 2DS version had higher average scores than all of the 3D versions. However, the gap between the mouse-controlled 3D displays and the conventional 2D1 display is smaller than in Experiment 1. In addition, without the mouse interface the two new 3D displays perform as well or better than the 2D1 version. Furthermore, the 3DP (programmed viewpoint) version was the best of the six in all but one of the non-collision scenarios (Figure 7).

Compared to Experiment 1, remarkably different results were obtained regarding accuracy. Overall, in Experiment 2 the 2DS version yielded consistently more accurate predictions than the other versions (Table 1). However, accuracy in the 3DC (cockpit-centered) and 3DP versions was as good as or better than that of the 2DS version when there was no collision (false alarm). One possible explanation for this is the chiefly horizontal view angle used in 3DC and 3DP, highlighting the differences in altitude between the aircraft. Also in contrast to Experiment 1 is the decreased accuracy of the 2D1 version, which might be accounted for by the increased mental workload of performing three-way numerical comparisons.

Table 1 Accuracy for Experiment 2. FA stands for false alarm percentage. FD stands for failure to detect collision percentage. PI stands for percentage incorrect overall. Note that lower values indicate greater accuracy.

	2D1	2DS	3DV	3DB	3DC	3DP
FA	43.2	22.7	40.9	38.6	22.7	18.2
FD	38.2	23.6	38.2	32.7	38.2	36.4
PI	40.4	23.2	39.4	35.4	31.3	29.4

Unlike Experiment 1, the 2D1 version in Experiment 2 had mean scores that were equivalent to all of the 3D versions at a 95% significance level. This supports the hypothesis that 3D is more helpful when the task is more difficult. The 2DS version had statistically higher mean scores at a 95% significance level than all of the 3D versions except for 3DP which it outperformed at only a 90% significance level. Based on these results, it appears that eliminating the need to find good viewpoints in the 3DC and 3DP versions allowed subjects to make faster, more accurate decisions than in the mouse-controlled 3D versions (3DV and 3DB).

Post-trial comments indicated that most subjects continued to find the 2DS (split-screen) version effective, but matching up aircraft with their counterparts on the other screen was more difficult with three aircraft. Among the 3D versions, subjects tended to prefer the 3DP version. In general, they still disliked having to use the mouse to locate a good viewpoint in the 3DV and 3DB versions.

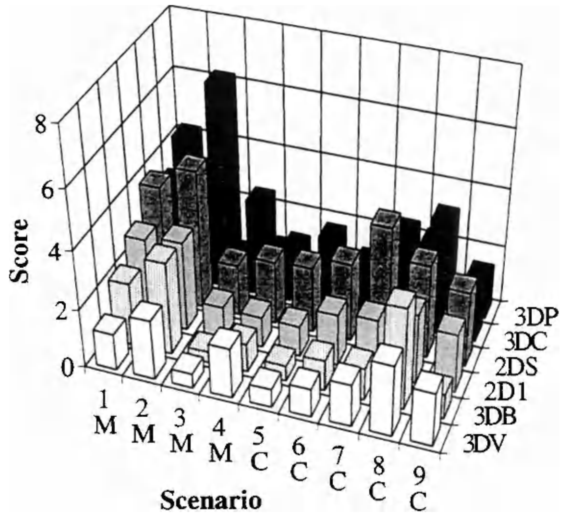


Fig. 7. Average scores for each scenario by version in Experiment 2. 'M' denotes a miss; 'C' denotes a collision.

6. FINAL EXPERIMENTS AND RESULTS

To further explore the idea that performance in the 3D versions might be approaching that of the 2DS (split-screen) version, some informal follow-up experiments were done. Six scenarios were created for the 2D1, 2DS, and 3DP versions with four and five aircraft. From the reactions and the scores of the few subjects that were run, it appeared that the 2DS version was still the most effective even with five aircraft.

7. CONCLUDING REMARKS

For simple straight-path scenarios involving two aircraft, the experiments indicate that collision prediction is best aided by a multi-viewpoint 2D display like the 2DS version above. The 2DS version filters out much of the irrelevant information, presenting all three dimensions of the aircraft's locations in a schematic, graphical manner using familiar overhead and horizontal views. However, in some cases a well-designed 3D display can outperform the conventional 2D numerical altitude displays and perform competitively with the 2D split-screen version, particularly in more complex scenarios.

Predicting collisions is only a small part of an air traffic controller's job. Further work needs to be done to determine if any of these new displays can help improve a controller's situational awareness or performance in more complex tasks such as runway approaches. On the other hand, because the experiments described here compared 2D schematic displays with perspective views of 3D scenes for tasks of varying visual complexity, the results may be useful for the more general problem of data visualization, an area that is drawing increasing attention from a variety of fields (Wickens *et al.*, 1994).

Whatever the application, the assumption that a "good" display must offer a high degree of "presence" is not always justified. The display that removes extraneous information, emphasizing only that which is essential to the task, is bound to win in the long run.

8. ACKNOWLEDGEMENTS

The authors would like to thank Profs. John Hansman and Tom Sheridan for their guidance and encouragement during the early stage of this project, Prof. Al Drake for his statistical guidance, and the colleagues from the Sensory Communications Group and the MIT community who graciously volunteered their time and constructive advice as participants in the experiments.

9. REFERENCES

Barfield, W. and Y. Kim (1991). Effect of geometric parameters of perspective on judgments of spatial information. *Perceptual and Motor Skills* **73**(2), 619–623.

Das, H. (1989). Kinematic control and visual display of redundant teleoperators. PhD thesis. Massachusetts Institute of Technology. Dept. of Mechanical Engineering.

Durfee, W. K. and Jr. R. M. Willis (1992). Simple depth cues for representing 3-d workspaces on 2-d computer displays. *IEEE Journal of Robotics and Automation.*

Harman, W. H. (1989). TCAS: a system for preventing midair collisions. *The Lincoln Laboratory Journal* **2**(3), 437–456.

Levison, W. H., R. W. Pew and D. J. Getty (1994). Application of virtual environments to the training of naval personnel. Technical Report Report No. 7988 (Draft). BBN Systems and Technologies. Cambridge, MA.

McGreevy, M. W. and S. R. Ellis (1986). The effect of perspective geometry on judged direction in spatial information instruments. *Human Factors* **28**, 439–456.

Wickens, C. D., D.H. Merwin and E. L. Lin (1994). Implications of graphics enhancements for the visualization of scientific data: dimensional integrity, stereopsis, motion, and mesh. *Human Factors* **36**(1), 44–61.

CONTROLLER-HUMAN INTERFACE DESIGN FOR THE FINAL APPROACH SPACING TOOL*

Maria C. Picardi

MIT Lincoln Laboratory, Lexington, Massachusetts

**This work was sponsored by the Federal Aviation Administration.*

ABSTRACT – The Federal Aviation Administration is developing a set of software tools, known as the Center-TRACON Automation System (CTAS), to assist air traffic controllers in their management and control tasks. CTAS originated at National Aeronautics and Space Administration (NASA) Ames Research Center, where prototypes continue to evolve. In parallel, Massachusetts Institute of Technology/ Lincoln Laboratory (MIT/LL) is refining and testing the software, including the Computer-Human Interface (CHI). This paper focuses on the CHI designed by MIT/LL for the Final Approach Spacing Tool (FAST) part of CTAS. The FAST design approach, CHI development and operational concept is presented.

Keywords: air traffic control; automation; human-centered design; human factors; human-machine interface.

1. INTRODUCTION

FAST comprises new computer automation software designed to assist radar controller teams in efficiently sequencing flight arrivals and precisely spacing aircraft on final approach during busy periods at terminals. Automation advisories generated by FAST are computed with respect to such real-world uncertainties as winds and weather conditions, pilot response times, aircraft performance, navigation accuracy, and communication delays. The FAST advisories include recommended sequence and runway assignment, accompanied by intelligently timed speed reductions and turns to specific vectors. They are displayed to controllers on Full Digital Arts Displays (FDADs). Whether or not the controllers can effectively interact with sophisticated automation aids presented to them on these older-generation, monochrome displays is a key human factors issue.

The FDAD workstations currently employed in terminal radar control rooms limit the Computer-Human Interface (CHI) design options. While some cryptic graphics are displayed on the FDADs (small arcs for vectors and splats for speed reduction points), most of the advisory information has been added as alphanumeric augmentation to the existing data blocks presented near each aircraft's radar symbol. Currently, only critical and timely information is displayed on the data blocks, to aid the controllers in their primary task of maintaining separation. The CHI design challenge was to augment the data block with an operationally suitable presentation of the advisories. Such human factors issues as target recognition, decision making, spatial perception, memory, use of cognitive maps, and minimal procedure modifications were considered in designing the output and associated inputs for the FDAD implementation of this interactive tool.

1.1 FAST CHI Output

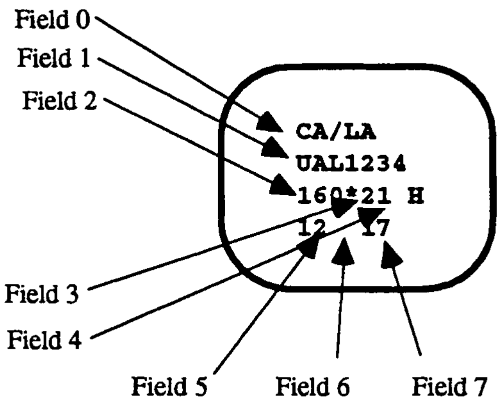
The system used in Terminal Radar Approach Control rooms (TRACONs) today, known as Automated Radar Terminal System (ARTS), displays full data blocks comprised of five data fields presented on three lines. In consecutive order, these fields contain the conflict alert and/or low altitude alert indicator (usually not visible), the aircraft identifier, the Mode C altitude or runway number, status information (an asterisk indicates controller entry in Mode C field, which timeshares with the Mode C readout), and ground speed or the aircraft type and the aircraft's weight class indicator (e.g., H for Heavy). Our design proposal is to present FAST data on an additional fourth line comprised of three more fields, as depicted in Figure 1. We refer to this technique as data block augmentation.

In the FAST CHI design, the added fields are sized and reserved for mutually exclusive advice (see Table 1). Certain advisories are mutually exclusive, that is, they do not time share with any other advisories. Rather, they exclude the display of each other as well as other possible advisories for a given field. This is important for precluding a cluttered look as data blocks typically surround one another on a busy radar screen and can overlap.

For example, the following advisories are mutually exclusive:

- Sequence number and resequence number, since the former is absolute and the latter is a relative number that is displayed briefly while the logic works on the request to change sequence.
- Controller discretion and the speed or heading number, since controller discretion logic inhibits generation of speed and turn advisories.

• Early/late indication, delay indication, and advice non-computable indication, since early/late is measured against the nominal route, delay is measured against the slowest route given the slowest time, and FAST is unable to compute advice when a trajectory error has been returned from CTAS.



Contents of fields (ARTS uses 0 through 4, FAST uses 5 through 7):

Field 0 - Conflict Alert, Low Altitude Alert

Field 1 - Aircraft Identifier (ACID)

Field 2 - Mode C Altitude/Scratchpad timeshared

Field 3 - Status

Field 4 - Ground Speed in knots and Heavy Indicator (if applicable)/Aircraft Type timeshared

Field 5 - Sequence Number/FAST Runway Indication possibly timeshared or Resequence Number or Missed Approach Advisory

Field 6 - Blank space

Field 7 - Non-Computable Indicator or Priority Aircraft Advisory or Controller Discretion Symbol or Delay Indicator or Speed Advisory Number/STAR Turn Advisory Number possibly timeshared

Figure 1. Existing ARTS data block augmented with FAST advisories.

Judicious use of time sharing succinct bits of information is also used for displaying some of the FAST augmentation to the data block. Controllers have become accustomed to viewing information that is continuously updated but abbreviated in both presentation and length of time interval. In the existing ARTS, three of the full data block fields are capable of time sharing displayed information. These fields can time share independent of one another. The unit time is 1/2 second. The number of 1/2 second intervals for the duration of each field is controllable via software. ARTS time shares fields 2 and 4 in the third line of the data block while FAST will time share augmentation fields 5 and 7 in the fourth line. FAST displays advisories at the same rate as the vertically aligned ARTS data, although different rates are technically possible.

If some advisories are not available at a given time, then the corresponding reserved data fields will remain blank. This is consistent with display of the ARTS' fields directly above, where information is presented in reserved spaces. Representative samples of augmented data blocks are presented in Figure 2.

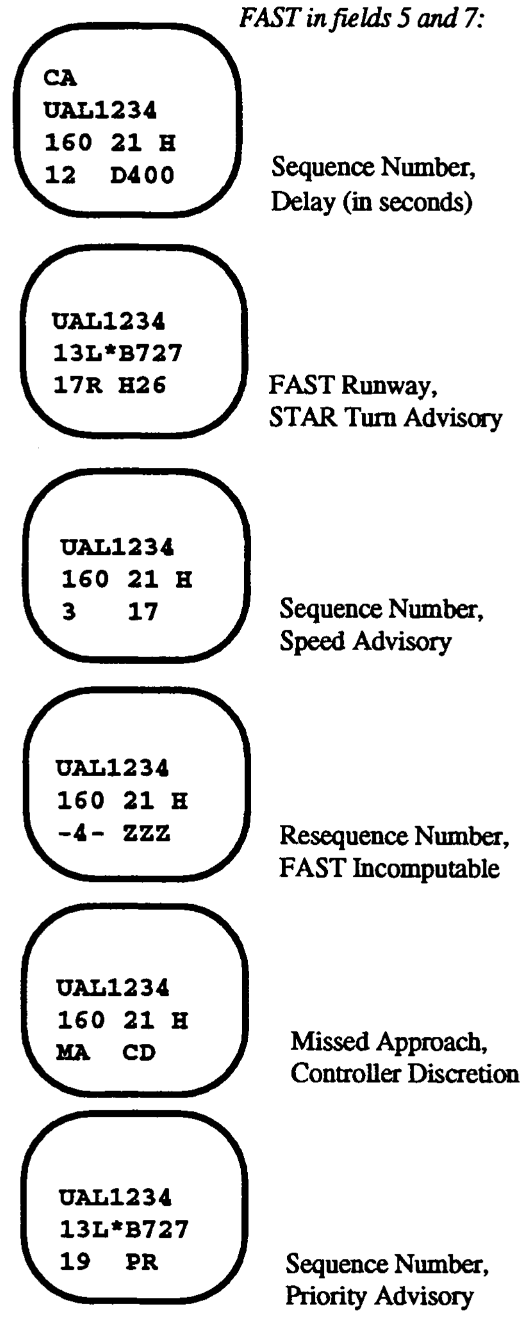


Figure 2. Sampler of FAST data block fields with advisories.

Table 1 ARTS data block augmentation fields of FAST advisories

Field #	# of Characters	Sample Content	Meaning
5	3	-4-	Resequence Number or
		2	Sequence Number or
		MA	Missed Approach Advisory
		13L	*possibly time shares with*
			FAST Runway Indication
6	1	*blank space*	none
7	6	ZZZ	FAST advice incomputable or
		CD	Controller Discretion or
		D400	Delay (in seconds) or
		PR	Priority Aircraft Advisory
		17	Speed Advisory Number
			time shares with
		H26	STAR Turn Advisory Number

CLEAR	BACK SPACE								ENTER
TRK START	TRK REPOS	TRK SUSP	TRK DROP	HND OFF	FLT DATA	MULTI FUNC	F8	Δ	.
F9	F10	F11	F12	F13	F14	F15	F16	IFR +	VFR /
A	BCN B	CFG C	DIS D	EMG E	FIL F	G	1	2	3
H	I	J	K	LDR L	MOD M	N	4	5	6
OFF O	PRE P	Q	R	SYS S	TAB T	U	7	8	9
V	W	X	Y	Z	*	∀	◊	0	□

Figure 3. ARTSIII Keyboard Layout Diagram

1.2 FAST CHI Input

As discussed above, the amount of output from FAST onto the FDAD – already crowded with such critical information as radar targets, data blocks, range rings, weather areas, satellite airport symbols, and tabular lists – was intentionally kept clear and concise. The same guiding principal holds for the additional input needed to support controllers' communications with FAST. Although the human factors design goal was to build a system as transparent as possible to the user, some essential FAST-specific keyboard entries were necessarily added as part of the CHI. In keeping with the development approach of introducing automation that is unobtrusive and easy to use, we determined that the existing ARTS keyboard (see Figure 3) and trackball be used for FAST input, which is consistent with our recommendation of augmenting the existing display for FAST output. The CHI design for FAST related input/output, therefore, should result in minimal impact or change to the controller's current operational concept.

Specifically, a currently unused function key (F12), and some unused shape keys (inverted triangle, square), were chosen to initiate the FAST inputs and differentiate them from the existing ARTS inputs. Within ARTS, certain keys have over the years been patched into the software to abbreviate frequently used functions. (Some of these are actually labelled on the keys and the same could be done for FAST inputs once acceptance is determined). For FAST inputs the number of keystrokes was kept down to one or two wherever technically feasible, with the exception of turning the transmission of ARTS data to FAST on and off since inadvertent activation is not desirable. Each input is echoed (i.e.,0 printed out verbatim as it is typed) in the ARTS preview area, already established for the purpose of echoing ARTS related inputs. The limited set of editing features in ARTS (clear, backspace) work consistently with FAST inputs. A separate small area called the FAST information list (FIL) is displayed to indicate the status of FAST (a simple rotating arrow reflects the system is alive), provide feedback for FAST-specific inputs and relate system messages.

The ENTER key on the ARTS keyboard has the same effect as using the ENTER key attached to the trackball (which is mounted near the keyboard on the FDAD console's shelf). As with the existing ARTS, either method may be used for most inputs. Table 2 presents the FAST-specific inputs currently supported by the software. The abbreviation "slew" refers to the action of moving the cursor (with the trackball) to cover the control position symbol of the desired data block or to pinpoint a spatial location on the FDAD. The acronym "ACID" is used to indicate the aircraft identification number may be keyed in. Where two input methods are listed for a given function, as in the ARTS either method will work. The FAST-specific output resulting from these inputs can be "quick looked" which is an ARTS feature that allows a neighboring controller in the same facility to temporarily view data blocks under the control of other radar positions. These outputs that are supported by the quick look feature are marked with an asterisk in Table 2.

1.3 Progression of CHI design and development

The CHI for FAST, indeed for all of CTAS, is being designed and developed in an iterative fashion with respect to the overall project philosophy of "human-centered" automation. Many different controllers from each major region in the United States, as well as some from other nations, have been exposed to FAST on the monochrome FDADs (and on color SUN workstations that will not be delivered into the field). The FAA has supported this design effort by providing System Development Teams (SDTs) who tested the early prototypes in the laboratories at MIT/LL and NASA Ames (Davis, et al., 1990). Additional controllers continue to experiment with FAST in a non-operational mode at the field development sites of Denver and Dallas/Fort Worth (DFW) where Cadres of their peers are trained to teach them the tool. Currently, a FAST Assessment Team is conducting a final review of "passive" FAST, i.e., sequence numbers and runway assignments. Formal training of all controllers at DFW will follow.

Close contact with the users of this product has allowed us to propose a design that was conceived and tested with human factors in mind. The first test of FAST on FDADs was conducted by MIT/LL using static symbology as part of a risk reduction series geared toward discovering if the tool could successfully migrate from a laboratory platform to a display used in the "real-world" of ATC. Up to this point, the SDT had seen FAST only on smaller color raster displays. Results from this risk reduction test, done at the FAA Technical Center (FAATC) confirmed that the CHI of the FAST data block augmentation was readily acceptable to the SDT (Picardi, 1992). Responses revealed strong agreement that the added FAST symbology does indeed appear distinct from the existing ARTS and videomap symbology. Therefore, confidence was gained that the passive advisories are being presented in an operationally suitable manner.

For further investigation of the CHI and other engineering aspects, increased simulation fidelity was needed. Therefore, a FAST System Test Environment was conceived and built by MIT/LL (Spencer, 1993). We used the FSTE to conduct a shake-down test of FAST with dynamic simulations using realistically heavy traffic scenarios and highly experienced full performance level controllers from Level V (the busiest) facilities The results from this dynamic real-time simulation where consistent with the SDTs initial favorable reaction to the FAST CHI. Based on the increased information displayed for FAST outputs, and timeliness required for FAST inputs, the controllers suggested limiting the number of advisories able to be quick looked and abbreviating the longer keystrokes (e.g., resequence). A copy of the FSTE was then delivered to NASA Ames to support their continued use of real-time simulations with the SDT (Picardi, 1992), and to FAATC where functionality demonstrations are conducted before each incremental field deployment to the development sites.

One key finding during tests conducted on the FDADs has been that the active advisories, which included the display of turn arcs and speed splats, are not as discernable to the SDT and they have commented on their perception of screen clutter, with suggestions for alleviating it. Based on the SDTs suggestion resulting from the potential for confusing the turn advisory with other ARTS symbology, the vector number presented adjacent to the detached arc was removed and placed only in the data block.

Table 2 FAST-specific inputs supported on the FDAD keyboard

Entry Format	Function
F12, ON, Enter	This enables transmission of ARTS data to FAST.
F12, OFF, Enter	This disables transmission of ARTS data to FAST.
F12, E, Enter	This enables the display of FAST advisories at all positions. The FIL rotating indicator is shown when advisories are enabled.
F12, I, Enter	This disables the display of FAST advisories at all positions, but ARTS data is sent to FAST. The FIL rotating indicator is not displayed.
F12, O, Enter	This toggles FAST advisories on/off at the entering position. When FAST advisories are off, the letter O appears in the FIL as a status indicator. The FIL rotating indicator is not affected by this entry.
F12, ⊟, Enter	This toggles turn and speed advisories on/off at the entering position. Other FAST advisories/displays are not affected. When turn and speed advisories are off, the character ⊟ appears in the FIL as a status indicator.
F12, X, ACID, Enter (or Slew)	This toggles scheduling status of specified aircraft between suspended and scheduled. These changes in scheduling status must be acknowledged at the traffic management display. When an aircraft is suspended from the schedule, the indicator X replaces the sequence number in the data block.*
F12, F, Enter	This clears the FIL display of any error messages or resequence messages.
F12, F, Slew	This relocates the FAST Information List.
F12, Δ, ACID, Enter (or Slew)	This toggles the priority scheduling status of the aircraft. When an aircraft is being given priority scheduling, the indicator PR appears in its data block augmentation.*
F12, •, ACID, Enter (or Slew)	This designates the flight as a missed approach to FAST. The indicator MA will appear in the data block augmentation, replacing the landing sequence number. FAST may display MA automatically as well.*
F12, A, ACID, Enter (or Slew)	This causes FAST to schedule a missed approach aircraft.
Rwy, Slew	This assigns a landing runway to the aircraft. The ARTS initially assigns a default landing runway to DFW arrivals based upon the entry fix, and displays this runway in the scratchpad area of the data block. If FAST agrees with this runway assignment, it will not issue any runway advisory. If FAST recommends a different runway, it will display a runway advisory in the data block augmentation. The controller only needs to make this keyboard entry if the default runway assigned by ARTS is unsatisfactory, or if FAST advises a different runway and a choice must be made. FAST will accept any runway assigned by the controller. Rwy is a single keystroke. For DFW arrivals the keys are as follows: + for **13R**, / for **31R**, • for **17R/35L**, Δ for **18L/36R**, ◊ for **17L/35R**, ∇ for **18R/36L***
F12, 5, ACID, Enter (or Slew)	The identified aircraft is given the landing sequence number 1, and all subsequent aircraft assigned to the same runway are renumbered accordingly.
F12, 5, Rwy, Enter	All aircraft assigned to the indicated runway have their landing sequence renumbered starting with 1. Rwy is a single key as for the runway assignment entry.
F12, R, ACID, Enter (or Slew)	This resequences aircraft. The aircraft are entered in the order of their new sequence. The first aircraft entered retains its current landing sequence number, and the other aircraft are sequenced behind it in the order they are entered. All entered aircraft must be assigned to the same runway. An aircraft can be removed from the sequence by entering it again, even though other aircraft have been entered in the meantime. While the aircraft are being resequenced, their sequence number is replaced by their relative sequence number in the format -n-. To facilitate the multiple entries, the ARTS multiple entry feature may be used by prefixing the first F12-R with M. This eliminates the need to enter F12 for subsequent entries, but the R must still be entered.

Currently, detached arcs are presented a short distance in front of the associated data block along the predicted flight path for the aircraft that will be issued a turn advisory. In laboratory testing on the development color workstations, both the arc and the data block are briefly painted in a unique distinct color to indicate their relationship. We suspect the lack of color to associate these detached advisories with their respective data blocks nearby on the monochrome FDAD and/or simply to call attention to them, i.e., situational awareness, is being reflected in the controllers' reaction.

A movement is underway in the FAA, known as the Stand Alone Replacement System (STARS), to deploy color consoles into TRACONs which may address one issue of acceptability when using active FAST advisories. We will examine the CHI design issues such as this one as the product continues to be re-engineered in preparation for eventual field deployment. MIT/LL recently began the process of Development Test and Evaluation (DT&E) on the re-engineered software for the Demonstration and Validation (D&V) phase of the project, during which the CHI of FAST and other parts of CTAS will eventually be formally verified. Meanwhile CHI design and requirements specification is undergoing final analysis and documentation by a team of researchers and users.

2. CONCLUSION

FAST, including the CHI described here, has been tested extensively by controller teams in laboratories at MIT/LL and NASA Ames and is currently being evaluated for use by controllers in Dallas/Fort Worth and Denver terminals. Technically, FAST data presentation and interaction has been shown not to interfere with the presentation or operation of the primary system, ARTS, as demonstrated at the FAA Technical Center (FAATC) during "noninterference" tests. Important human factors issues regarding the FAST CHI, such as workload, distribution of automation access between controller teams and traffic management coordinators, and maintenance of situational awareness, remain to be fully determined and addressed when the software is deployed for ATC operations. In the future, the potential use of data link between the pilot in the cockpit and the radar controller in the TRACON with respect to ATC automation will be an interesting challenge to consider. For now, early research and development experience accompanied by simulation testing results has indicated that the CHI successfully communicates the FAST concept to the user community and that, in turn, the concept appears a viable one for automating a carefully chosen portion of the air traffic control task.

In summary, the FAST CHI design retains consistency in appearance and behavior of data block items, while minimizing clutter and distraction. To enhance usability, spatial cues are provided for advisory placement on a single added line in the data block. Data presentation is abbreviated and presented to the controller only when timely. Controller input back to FAST was also designed to be minimal and consistent with their existing ARTS keyboard and trackball method. Progressively more robust testing with FAA-supplied controller teams has consistently shown that the CHI design for at least the first increment of this new automation (known as passive FAST) will be acceptable for operational use by air traffic controllers at the major TRACON facilities throughout the United States.

REFERENCES

Davis, T., H. Erzberger, S. Green. (May 1990) *Simulator Evaluation of the Final Approach Spacing Tool,* NASA TM-102807.

Picardi, M. C. (1992). *Risk Reduction Test Number 1 Report: FDAD Static Symbology Demonstration of FAST*, 41PM-TATCA-0006, MIT Lincoln Laboratory, Lexington, MA.

Spencer, D. A. (1993) *Interface Requirements Specifi-cation for the FAST System Test Environment, 41PM-TATCA-0009,* MIT Lincoln Laboratory, Lexington, MA.

INTEGRATING THE WORK FORCE INTO THE DESIGN OF PRODUCTION SYSTEMS

Frank Emspak, Assistant Professor

School for Workers, Department of Labor Education
University of Wisconsin-Extension, Continuing Education Extension
610 Langdon Street, 422 Lowell Hall
Madison, Wisconsin 53703

This paper will describe a project undertaken by the custom woodworking industry owners and skilled workers to identify criteria for the design of new software for their industry and attempts to have that software produced. Key criteria were skill and flexibility. The role of the NIST as a provider of support will be discussed.

Design systems; flexible manufacturing systems; man/machine interaction; skill based production.

1. INTRODUCTION

The paper describes a project undertaken by the School for Workers and the Technology Sub-Committee of the National Labor-Management Committee of the Custom Woodworking Industry (NLCCWI). The committe is one of numerous efforts in many branches of industry wherein skilled workers and management attempt to design new systems of production. The design focus contrasts with more widespread attempts to jointly implement existing production technology. Should such design efforts succeed, significant increases in productivity would be expected along with changes in job content which emphasize skill and human development.

2. BACKGROUND

The modern mass production system in the United States imposed an organization of work which has increasingly separated the design of tools from their users. Dramatic examples of the separation have been the development of computer-based tools of all sorts. The basic mathematical theories, the inventions of the computer chip and the computer, as well as the marrying of computers to machinery, was done primarily in laboratories and research institutes far from the point of use. The control of machine tools, once the province of the skilled machinist was also increasingly removed from the shop floor. Today, CNC machine tools are controlled by programs usually produced off the shop floor by specialists in programming. The programs are subject to after the fact changes if there is a problem. In particular the removal of planning and programming from the shop floor to the office has not been forced by technology but by the politics of control of the workplace which in turn was supposed to yield higher quality, higher productivity and thus lower costs (Brodner, 1985).

However, there is a growing realization among engineers and industry that designs which treat workers as an after thought are more and more problematical (Brodner, 1988; Schultz-Wild et al., 1985).

At the same time, various worker-led institutions have begun to understand that technology can be designed and implemented in ways that might preserve or enhance skill while at the same time meting more traditional business needs (IAM, 1994; AFL-CIO, 1994).

The humanization of work movements in Sweden and

Germany were the first widespread flowering of the notion that technology could integrated with the needs of workers. Humanization- that is less noise, safe working conditions, a human pace, more job variability, improved man machine interfaces, all take as a given the basic design criteria of the technology. A considerable step beyond the humanization of work has been the "socio-technical" school of design which attempts to bring more equality between technical criteria and social criteria. As a group of professionals socio-technical designers have argued for (and designed) holistic work organizations and more accessibility to programming for shop floor workers. Increased job flexibility, group work and teams, along with increased skills training are part of a socio-technical approach. Here in the US "socio-technical" design has been implemented in places such as the Saturn facility in Spring Hill, Tennessee, among other locations.

More recently a number of economic and social forces have emerged that places a redefinition of the man-machine and man-technology relationship at the forefront of design challenges. Among those forces are: 1) increasingly short product cycles; 2) a demand for ever higher quality standards; 3) segmented markets requiring more flexibility in production; 4) the tremendous cost of designing systems; 5) the increasing use of JIT production models; and 6) pressure from the workforce itself.

3. COLLABORATIVE DESIGN

The issue that must be posed today goes beyond the notion that engineers should design systems that give weight to the users needs (socio-technical). The challenge today is to enable organizations to liberate the intellectual potential of all members so as to function better. In other words how can users (workers, suppliers and customers) and engineers work collectively to design systems that work for both the users and the people who commission the work?[1]

Changes both in theories of management and in perceptions held by working people regarding their role in determining work organization are taking place simultaneously. At the moment pressure from below is limited in scope but it is significant. The new view from below suggests that managing the firm-especially when it comes to the design of production is too important a decision to be left to management or engineers alone.

A reflection of these changes is evident in the existence of various forms of employee involvement activities in all manner of enterprises. Strong labor-management organizations exist in a variety of unionized industries. Examples of locations which have joint committees dealing with the **design** of the technology as a focus of their work include certain plants of the Ford Motor Company, all the basic steel industry and firms such as the Bath Iron Works. The basic steel companies have signed a contract with the United Steel Workers of America which specifically includes decisions regarding technology as an area for joint decision making (United Steelworkers of America). The Machinists union, as noted earlier, has also identified technology as an area for potential partnership with industry (IAM, 1994; McCubbin et al., 1994). In all three instances the unions and firms have carried their discussions beyond implementing existing systems and into the realm of the redesign of work and tentatively into the area of technology design itself.

4. CASE STUDY: THE CUSTOM WOOD WORKING INDUSTRY

In 1992 the possibility of an industry group working with labor to define the technologies that best suited their industry presented itself in the custom woodworking industry. The United Brotherhood of Carpenters and Joiners and a number of custom woodworking shop owners established the National Labor-Management Committee for the Custom Woodworking Industry (The Federal Mediation and Conciliation Service has supported the work of the NLMCCWI). The overriding concern was expressed this way by members of the committee "if the custom wood working industry is not successful in meeting the challenge ...of the production shops... it will be priced out of the market" (Grant, 1994).

The objectives of the committee were to define common interests, such as training, and to work out means to advocate collectively for those interests. In early 1992 the National Labor Management Committee established a technology sub-committee. The sub-committee set out to assess the technical base of the industry and to identify, if possible, the technologies best suited to the industry. Additional objectives included encouraging the design and use of suitable new technologies and integrating the new productive methods into the existing apprenticeship system. The technology sub-committee operates by consensus so decisions reached reflect commitment from all participants.

4.1 The Industry

The custom wood working industry is in the

[1] "User" is defined as blue and white collar workers as well as the customer. In a sense the worker in a production system is the ultimate customer.

secondary wood products sector. The industry includes one of a kind units and limited production runs of furniture, store fixtures, millwork (doors, windows, moldings), veneer paneling, casework (kitchen cabinets, custom office cabinets), and exhibits. All types of woods are used- solid and veneers, as well as plastic, particle board, and high and low pressure plastic laminate. As wood is a living substance and its composition changes with the species, temperature and humidity, shaping and finishing wood requires a high degree of judgment (Grant, 1994).

The industry is dominated by small family owned enterprises. Most firms employ less than 50 people. Many are located within urban areas and hence provide skilled urban based employment. The result is a fragmented industry, with little money for capital investment and little or no research capacity.(Malakoff et al.,1995)

4.2 The Union as an Integrator

The dispersed and small size of individual shops means that individual firms cannot train its needed supply of skilled labor nor can the custom woodworking industry alone dictate the technology. The fragmented nature of the industry allows the Carpenters Union to play a major role as integrator and organizer. The union organizes the training of skilled cabinet makers through the nationwide system of apprentice training. The most advanced apprenticeship sites also provide a point to introduce new technologies to the industry (The apprentice program is primarily focused on the construction end of the trade. Most cabinetry programs are housed in or are part of the construction apprenticeship training programs. There are about 35 cabinet sites nationwide. They are managed by a joint industry-union organization and financed by a cents per hour worked contribution).

4.3 THE TECHNICAL LEVEL

The technical base of the industry resembles that of the job shop in metal working about twenty years ago. A survey commissioned by the carpenters union indicated that the penetration of CNC or NC equipment, CAD/CAM and even CAD was at a relatively low level. For example only 15% of the respondents reported working in shops with CNC routers, 13% in shops with CNC boring machines. About 37% of the shops use CAD (Malakoff et al., 1995). If there is modern equipment- CNC for example, it is usually not integrated into the shop. The survey data was substantiated by many in-depth interviews within the industry with both owners and skilled carpenters as well as extensive discussions with the major suppliers of wood working equipment in the United States.

The pace of technological change has been uneven. In the late 70s and early 80s new forms of material altered what the carpenters actually worked with. In the late 80s the demand for high quality mass-produced items brought about the use of CNC woodworking equipment. However, this equipment generally bypassed the custom woodworking shops because of its expense, inflexibility, and the lack of workers skilled in the use of the new equipment. By the late 1980s manufacturers of wood working equipment began to address the need for suitably scaled equipment for the custom industry. But still most modern machines stood alone and were not integrated into the facility (Grant, 1994).

5. THE COLLABORATIVE DESIGN PROCESS

The technology committee began its work in the Spring of 1993. At first the group concentrated on training itself in group problem-solving techniques so as to enable it to function. Secondly, the committee began a systematic survey of the industry to determine its technical level. Thirdly, the committee organized a series of visits and study tours to best practice firms, as determined by the committee. As the committee consisted of several skilled cabinet makers with long years of experience as well as owners from some of the most technologically advanced shops collectively the committee rapidly gained a mutual understanding of the state of the industry.

The successful training phase enabled all members of the committee to reach a common understanding of the industry and for each member to understand and appreciate the needs of the other members. The committee also educated itself regarding specific types of software and machinery currently on the market. Once the committee felt comfortable with its knowledge the group turned its attention to reaching consensus on the character of industry they would like to see in the future. In this regard there was considerable agreement. First, the committee decided the industry should be based on the use of skilled labor trained in a modern apprentice system. Secondly, the group determined that the needs for flexibility and customer satisfaction as well as quality could be met by small shops providing those shops modernized and made use of the potentials for the new technologies. We note that similar concerns regarding the holistic nature of work and the appropriateness of automation in flexible settings have been raised elsewhere (Martin et al., 1990; Martin, 1990).

5.1 Design Criteria

Given these values the committee then undertook to reach consensus on the criteria for the technologies that would be of greatest utility. The sub-committee adopted nine criteria. They wanted:

- machinery to be flexible - defined as capable of being used in production of one of a kind items;
- a technology system that would encourage self-learning, which is defined as cabinetmakers, drawing on their craft skills and way of conceiving work, able to utilize to use machinery, perferably without learning a whole programming language;
- systems that would improve competitiveness of the firms;
- equipment or systems that could decrease product throughput time;
- ergonomic considerations such as noise, fumes and repetitive motion syndrome to be dealt with in the design of the equipment;
- enhancing the skill of the craftsman "...the committee doesn't...want machinery that reduces the skill content of craftman's jobs..the jobs of workers using machinery should not be made dull and repetitive" (Carpenters, 1994);
- "simple, elegant" machines that skilled workers can use exercising their craft judgement;
- affordability (given the small enterprise nature of the industry) - machines should be priced at $100,000 maximum with payment over three or four years;
- machines that were durable and had low maintenence cost.

The design criteria are to be given equal weight should a software system or machines be produced. In many respects the criteria developed by the committee mirrored the theoretical discussions that had taken place abroad in the late 1980s (Hirsch-Kreinsen et al., 1989; Corbett, 1987).

Practically speaking the criteria meant that equipment had to be quieter but also be able to cut and shape rapidly. In terms of CNC/NC equipment the software had to be designed in such a fashion as to encourage shop floor programming and to build on the tacit knowledge of the carpenter- especially as it related to the changing and variable nature of wood. This notion of building on the knowledge of the carpenter and designing controls and software to do that has two aspects. One suggests using the knowledge of the carpenter to assess the conditions of the material rather than building expensive sensing and complicated software. This idea follows from the design criteria of low cost and the agreement of the value of skilled work. The second aspect addresses the view that programming using existing software basically asks the skilled worker to learn a different language and conception of his/her work. The committee feels that it would be most helpful to develop software that enhances rather than replaces the thought process of the skilled carpenter.

Overall there was a desire to enhance a work organization that is basically collective in content, rather than breaking up the work into discrete packages- such a programming, design, etc. In other words the committee decided that the concept of "craft", that is the unification of conception and execution, needed to be enhanced because it is a positive good for the industry and, in fact, allows the custom woodworking industry to exist in the first place. As this is a custom business often the carpenter as well as the owner or salesperson is interacting with the customer and as a result "designing" the work he will do. In other words the craft concept is the dominant view that most workers have of themselves. In turn that notion of craft which is equated with quality is valued by the owners.

5.2 Next Steps

The next step in the process was to decide if it was possible to actually commission the production of machinery or software that would match the design criteria. Basically the committee felt that the existing machinery was close enough to the needs expressed by the committee to be utilized without significant design changes. The software was another story. Starting in the Autumn of 1993, the committee actively pursued means to find funds and to hire appropriate software designers who would work under the guidance of the committee.

6. THE NATIONAL INSTITUTE FOR STANDARDS AND TECHNOLOGY (NIST)

In the Summer of 1993, members of the committee met with representatives of the Advanced Technology Program (ATP) of the NIST in an attempt to find funding for the implementation phase of the project. Two related courses of action were followed. On the one hand the committee applied for funding in the "General Competitions" sponsored by the ATP. Secondly, the author worked with NIST to establish a focused program which would specifically fund collaborative design projects providing they met the other technical and business criteria of the Advanced Technology Program. In either case the objective was to achieve funding to implement a project in which the design criteria was the result of collaboration between skilled workers and engineers and which would result in design ideas superior to those derived from more traditional development practice (Emspak, 1994).

Unfortunately neither source of funding was forthcoming. NIST believes that the emphasis on the process of design, even including the possibility of developing computer-based design tools, is too sociological and does not meet the hard technology mission of the organization. At this point the committee has explored other sources of support both from government and industry. It is unclear if the committee will be able to commission the design of software to meet its standards.

7. CONCLUSIONS

The project of the Sub-Committee on Technology of the National Labor-Management Committee of the Custom Woodworking Industry was distinguished from other joint efforts in that it went beyond the collective implementation of a production system chosen by management. Nor did the committee simply try to choose between existing systems based on some evaluation of "user friendliness". Rather it was an attempt by owners and workers to collectively define a production system based on shared values and then actually commission its creation. Of significance in terms of the man-machine discussion was a realization by both management and labor of the importance of skill and the their subsequent ability to develop criteria which would incorporate skill in new production technologies. Thus from a planning and design point of view the committee's work was a success.

However, the committee was not able to produce the product it desired- not for lack of volition, but for lack of development funding. Perhaps this situation is mirrored in other industries of similar make up, and if so, it would suggest that if collaborative design is to move forward in this country we need to develop a mechanism to encourage it.

REFERENCES

AFL-CIO Evolution of Work Committee (1994). Policy Statement *The New American Workplace: The Labor Perspective*. AFL-CIO Washington.

Brodner,P. (1985). Qualification The Superior Choice to the *Unmanned Factory. 2nd IFAC/FIF/FORS/IEA Conference on Analysis, Design and Evaluation of Man Machine Systems*, Varese Italy, September 10-12, 1985

Brodner,P.(1988). Options for CIM: *Unmanned Factory* versus Skill-Based Manufacturing. *Computer Aided Manufacturing Systems*, Vol. 1, No. 2, pp.67-74.

Corbett, M. (1987). Strategic Options for Human- Machine Interface Design. In: *Strategic Options for "New Production Systems- CHIM: computer and Human Integrated Manufacturing* (P.Brodner, Ed.) pp. 157-174 FAST Occasional papers No. 150 p.

Emspak, F. (1994). Focused Program Recommendation: Integrating Workers in Design Engineering.(unpublished) Presentation to NIST-ATP.

Grant, M. (1994). Computerized Machinery and Work Reorganization in the Custom Woodworking Industry. Presentation to the NIST-ATP regional meeting, Seattle Washington October 4, 1994.

Hirsch-Kreinsen, H. and Schultz-Wild, R. (1989). Implementation Processes of New Technologies-Managerial Objectives and Interests. *Automatica* paper Number 87-29

International Association of Machinists and Aerospace Workers (IAM), (1994). High Performance Work Organization Partnership IAM, Washington D.C.

McCubbin,J.G.,and Pappenfuss,D. (1994). Involving the Workforce in the Design of New Process Equipment. Presentation to the NIST-ATP Regional Meeting Milwaukee Wisconsin, September 29, 1994.

Malakoff, W. and Kim, D. (1995). Computerized Technology, Jobs, and the Union's Role: The Views of Union Members. publication Summer 1995.

Martin,T., Ulich,E. and Warnecke, H.J., (1990). Appropriate Automation for Flexible Manufacturing. *Automatica*, Vol 26, No 3, pp 611-66.

Martin, T., (1990). The Need for Human Skills in Production - The Case of CIM. *Computers in Industry*, 14 pg 205-211, Elsevier Science Publishers B.V. North Holland

Schultz-Wild, R. and Kohler, C.(1985). Introducing New Manufacturing Technology: Manpower Problems and Policies. *Human Systems Management* 5 , pp. 231-243; Elsevier Science Publishers B.V. North Holland.

United Brotherhood of Carpenters and Joiners,(1994). Minutes Mill-Cabinet Technology Committee, January 11, 1994, Washington, D.C.

United Steelworkers of America,(1993). Memorandum of Understanding on Inland/USWA Partnership, May 21, 1993.

AN ADAPTIVE TROUBLESHOOTING MODEL FOR COMPLEX AND DYNAMIC SYSTEM

H. Jesse Huang

School of Industrial Engineering
Purdue University
W. Lafayette, IN 47907

Abstract: Gaining insights into the principles underlying human reasoning and troubleshooting strategies is likely to lead to a more effective and intelligent computer-assisted interface design in diagnosis task. In this paper, an adaptive neuro-fuzzy model, which explores the latter approach, is proposed conceptually to implant human reasoning abilities and exploit human problem solving processes to diagnose novel faults by incorporating the knowledge association of fuzzy logic into the neural networks structure. The adaptive system, a hybrid of neural networks and fuzzy logic inference, has the potential to capture the benefits of three levels of troubleshooting knowledge, low level statistical results, middle level designer's control knowledge, and high level operator's field technique, into a single capsule.

Keywords: adaptive system, fault diagnosis, fuzzy hybrid system, neural networks

1. INTRODUCTION

Gaining insights into the principles underlying human reasoning and troubleshooting strategies is likely to lead to a more effective and intelligent computer-assisted design in diagnosis task. Solving this problem has been, and remains, a most challenging and fascinating long-range goal in human-machine interaction design.

1.1 Review of novel diagnosis in complex systems

Traditionally, decision theoretic approach produces descriptions that are too mathematical and difficult to comprehend for a human expert [Rouse, 1984]. This even prevents the human user from intervening in the analysis with knowledge resources and observed failure symptoms. While expert systems have been built using large collections of rules based on empirical association or decision logic, interest (Fath *et al*, 1986; Govindaraj *et al*, 1986) has grown recently in the use of systems that provide deeper knowledge and human-like reasoning ability when solving problems that are either unfamiliar or unexpected to the human user. It has been reported (Yoon and Hammer, 1986; Fath *et al*, 1986) that the knowledge-based rules prevent human expert from using features which are known to be strongly dependent on each other, even if these features and relationships existing among them are actually by expert during the classification process.

Hart (1982) and Michie (1982) have been written about the depth of knowledge should be represented in problem-solving model that is able to incorporate a variety of relative knowledge. The distinction is that "surface systems" are at best a data base of pattern-decision pairs or condition-action rules. "Deep systems" will capture the human's pattern recognition ability and so called "first principles" human reasoning strategy.

Studying about the strategies for diagnosis, Rasmussen (1986) pointed out two diagnostic strategies-symptom strategy and topographic search-are generally applied by human expert in diagnostic tasks. While symptom strategy is using a set of observed symptoms to match a library of abnormal system conditions, topographic search is comparing a normal model in order to find the location of a failure. It has been argued that either one can't provide deep knowledge for complex reasoning. Duncan *et al* (1980) and Morris and Rouse (1985) have conducted similar experiments on trained strategies on fault diagnosis experiments. On novel faults, both research found operators with only theory knowledge are incapable to solve novel problems comparing to other two groups which were tained with either rules and nothing. A possible implication of these empirical experiments is that an aid should contain a system model with human feature recognition and "first principles" reasoning abilities with difficult situations.

1.2 Human aspects of problem solving

Anderson's ACT* (Adaptive control of thought) is an elaborate theory that accounts for a paradigm about the acquisition of problem-solving skills. The ACT* theory assumes a problem solver will try to solve a problem by analogy to previous knowledge resources of a similar example (Anderson, 1983). ACT* provides learning mechanisms regarding knowledge transition in a knowledge-rich domain. While Rasmussen (1986) states a diagnostic search aid should explore human's feature recognition ability and the functional description with expert's knowledge. A conceptual knowledge hierarchy paves the road for guiding user to solve novel problems.

2. NEURO-FUZZY MODEL FOR THE DIAGNOSTIC PROCESS

The recent development on incorporating the learning algorithm and nonlinearity clustering ability of neural networks into the human-like thinking process of fuzzy logic has been a great success applying in control algorithm and learning mechanism (Buckley *et al*, 1993; Carpenter *et al*, 1992). However, little studies have been done in simulated human cognition activities in the presence of imprecise data and uncertain knowledge by combining both benefits of neural networks and fuzzy system.

The combination of neural networks an fuzzy logic have three advantages as following.

(1) Cloning: transforms historical data set into patterns with similar solutions. This functionality reduces human's working memory load.

(2) Tracking: provides a learning trajectory to a reference reasoning. The diagnostic process from symptom patterns to system function leads user away from making incorrect hypothesis.

(3) Dynamic optimization: provides a reasoning tool to maximize utility based on different knowledge sources. The learning ability offers an important advantages over classical methods. The emphasized merit is the fuzzy rules has the ability of describing the problem solver's preference structure instead of the strict correspondence by traditional deductive rules (Whalen *et al*, 1982).

2.1 Mathematical background of neuro-fuzzy model

The mathematical background of Neuro-fuzzy model originates from fuzzy neuron concept which has been applied in a variety of domains (Gupta and Rao, 1994; Imasaki *et al*, 1992; Maesa and Murakami, 1993). In Figure 1,

X= input data (1)
ART1= adaptive reasoning theory (2)
Y_p^K= p clusters at k th step (3)
Z= output of neuro-fuzzy system (4)

First, the input pattern is compared with reference samples or clusters. Second, the computed reference similarities are translated into another domain by a membership functions. Third, to evaluate a total grade that input belongs to a certain class, then the membership grades of all reference samples in the class are combined by an AND operation. Finally, a defuzzification process is executed.

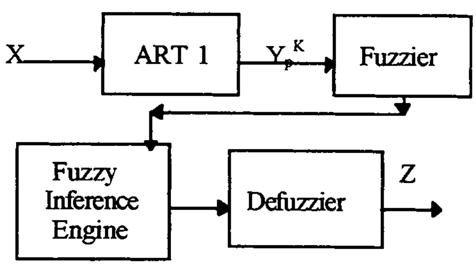


Figure 1 Block diagram of the neuro-fuzzy model

2.2 Computing reference similarities

As unsupervised model (when no human intervention is presented), ART1 performs clustering by comparing the minimum distance between each sample data vector X_i with the existing patterns **C**. The ART1 can accommodate a new cluster without affecting the storage of cluster already ;earned. Moreover, ART1 reflects the degree of similarity of those clusters. Euclidean distance is applied on computing reference similarity.

$$HD(X_i, Y_i) = \sqrt{(X_i - Y_i)^T (X_i - Y_i)} \quad (5)$$

$$X^t C^{(p)} = [n - HD(X, C^{(p)})] - HD(X, C^{(p)}) \quad (6)$$

$$net_p = \frac{1}{2} X^t C^{(p)} + \frac{1}{2} \quad (7)$$

$$net_p = n - HD(X, C^{(p)}) \quad (8)$$

2.3 Membership function

An entropy measure is introduced to represent the vagueness of fuzzy numbers, which is equivalent to determine the quantity of information contained in them. For finite reference knowledge sources.

$$K = \{k_1, k_2, ..., k_m\} \quad (9)$$

The membership degree for category A_i is defined by membership function as

$$\mu_{A_i}(k_j) : [0,1] \quad (10)$$

Mathematical operations are performed using logical operators, such as AND, OR and negation.

$$(X \vee Y)_i \equiv \max(X_i, Y_i) \quad (11)$$

$$(X \wedge Y)_i \equiv \min(X_i, Y_i) \quad (12)$$

$$\neg(X)_i \equiv neg(1 - X_i) \quad (13)$$

A linguistic table in Table 1 is used to determine linguistic value.

Table 1 Table for determining the linguistic values

	VLR	LR	MR	HR	VHR
VLE					
LE					
ME					
HE					
VHE					

VLE, LE, ME, HE, and VHE are utilized to describe the evidence degree for input vector X with each potential clusters, while VLR, LR, MR, HR, and VHR are used to describe the relationship for input X to different knowledge sources.

3. TASK EXPERIMENT

A forging press (Lin, et al, 1994) is applied in this task for subjects to diagnose and troubleshoot possible faults. Forging hydraulic machines are widely utilized as major manufacturing equipment in advanced industry. Figure 2 is the functional hierarchy of the simulated machine. The experimental data includes protocols and operator action sequences.

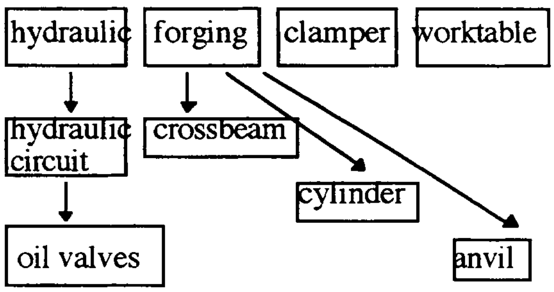


Figure 2 Hierarchy of simulated machine

4. CONCLUSION

Biology does provide a motivation and framework for the development of the Neuro-fuzzy structures. The approach of neuro-fuzzy models for diagnosing task represents an effort toward effectiveness and learning. The emphasis in this paper, both from mathematical structure and information processing point of view, is the confluence of measurement of the neural inputs and accumulated experiences in synaptic weights, and the linguistic fuzzy rules from designer's knowledge. With the development of the Neuro-fuzzy model, it is envisaged that learning schemes for diagnosing ill have following features:

(1) easy to implement fuzzy natural languages so that the structure of knowledge is clear and efficient.

(2) easy to accumulate user's experiences for any changed in the task and environment.

However, it should be noted that more research endeavors are necessary to develop general topology of the Neuro-fuzzy model, so that the system is made of applicable to model and control complex system.

References

Anderson, J. R. (1989). "A Theory of the Origins of Human Knowledge," *Artificial Intelligence*, 40, pp. 313-351.

Buckley, J. J., Y. Hayashi, and E. Czogala (1993). "On the equivalence of neural nets and fuzzy expert systems," *Fuzzy Sets and Systems*, 53, pp. 129-134.

Carpenter, G. A., S. Grossberg, N. Markuzon, J. H. Reynolds, and D. B., Rosen (1992). "Fuzzy ARTMAP: A neural network architectures for incremental supervised learning of analog multidimensional maps, " *IEEE Trans. Neural Networks*, **3**, #5, pp. 698-713.

Duncan, K. D., C. Cooper, and J. B. Brooke. (1983) "Effects of computer-aided and pre-training on fault location," *Ergonomics*, 26, pp. 669-696.

Fath, J. L., T. Govindaraj, and C. M. Mitchell (1986), "A methodology for development of a computer aided instruction program in complex, dynamic systems," *Proceedings of the 1986 International conference on systems, man, and cybernetics, Atlanta, GA*., pp. 1216-1221.

Govindaraj, T, and Y. D. Su (1988) "A model of fault diagnosis performance of expert marine engineers," *Int. J. Man-Machine Studies*, 29, pp. 1-29.

Gupta, M. M., and D. H. Rao (1994). "On the principles of fuzzy neural networks, " *Fuzzy Systems*, **vol. 1**, #1, pp. 171-183.

Hart, P. E. (1982). "Direction for AI in the eighties," *SIGART Newsletter*, **79**, 79.

Imasaki, N., J.-I. Kiji, and M. Arai. (1992). "Framework for Fuzzy Neural Networks," *SPIE*, **vol. 1710**, science of artificial neural networks, pp. 535-546.

Lin, H., Y. Yih, and G. Salvendy, "Neural-network based diagnosis of hydraulic forging presses in China," (submitted).

Maeda, H., S. Murakami (1993) "The use of a fuzzy decision-making method in a large-scale computer system choice problem, " *Fuzzy Sets and Systems*, 54, pp. 234-249.

Michie, D. (1982). "High-road and low-road programs, " *AI Magazine*, **3**, 1, pp. 21-22.

Morris, N. M., and W. B. Rouse. (1985) "The effects of type of knowledge upon human problem solving in a process control task," *IEEE Trans. Systems, Man ,and Cybernetics*, **vol. smc-15**, 6.

Rasmussen, J. (1986). **Information Processing and Human-Machine Interaction**, North-Holland, New York.

Rouse, W. B. (1984). Advances in man-machine systems research. (Ed.). JAI press inc.

Whalen, T., B. Schott, and F. Ganoe. (1982), "Fault Diagnosis in A Fuzzy Network," *IEE*E, pp. 35-39.

Yoon, W. C., and J. M. Hammer (1986). "using a qualitative model to aid the operator during novel fault diagnosis," *Proceedings of the 1986 International conference on systems, man, and cybernetics, Atlanta, GA*., pp. 1222-1227.

AN ORIGINAL "HUMAN-ORIENTED" ASSESSMENT APPROACH OF DESIGN METHODS FOR AUTOMATED MANUFACTURING SYSTEMS

J-C. Popieul, J-C. Angué

Laboratoire d'Automatique et de Mécanique Industrielles et Humaines
URA CNRS 1775
Université de Valenciennes, Le Mont Houy
BP 311, F-59304 VALENCIENNES Cedex, France
E-mail : popieul@univ-valenciennes.fr

Abstract: In this paper we first present a new method dedicated to the design of automated manufacturing systems, and then the evaluation of both its efficiency and its usability from the designer point of view. The originality of this work mainly lays in the evaluation approach which, rather than considering the sole technical aspects of the method (description power, formalism, ...), uses a real design context and aims at assessing the benefits gained by designers using this method rather than others.

Keywords: Manufacturing Systems, Design, Evaluation, Formalism.

INTRODUCTION

Today, industry, for its medium and wide range productions uses more and more automated means. These means have to meet quality, reliability, maintainability constraints that are more and more demanding. The satisfaction of these constraints and, as a consequence, the quality of the resulting system depends mainly on the decisions and choices made during the specification and design phases of the system. As a matter of fact, errors made during these phases will have significant consequences on the overall cost of the project (Smith 1988).

Regarding the importance of the specification and design steps joined to the increasing complexity and quality requirements of the designed systems, these steps can no more be realized without a methodological support. In this way, Automated Manufacturing Systems (AMS) designers, considering the lack of methods dedicated specifically to their work, use different methods extracted from other fields such as software engineering or consumer goods design.

A preliminary study allowed us to put forward that these methods were poorly suited to the use the AMS designers were making of them (Zaytoon, *et al.*, 1993; Lhoste, *et al.*, 1991; Popieul, 1994). In this way, we developed a new AMS specific design method which is named AMSD. This method is briefly described in the first part of this paper.

A first validation step of AMSD, consisting in using it on two actual contexts, allowed us to demonstrate the good representation capacity of the method. However, this kind of validation does not demonstrate its usability. To do so we have used an original validation approach based on the observation of designers in an actual AMS design problem context. This approach and the obtained results are detailed in parts two and three of this paper.

1. THE "AMSD" DESIGN METHOD

1.1 Proposed formalism

The proposed formalism is introduced figure 1. It is based on a hierarchic decomposition of the problem using a set of specific graphic symbols.

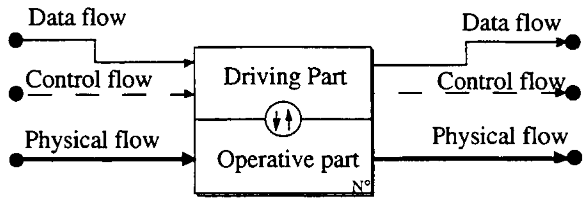


Figure 1: Proposed formalism for a diagram level different from zero

This formalism includes two main elements :

- "boxes" divided into two areas. The lower one symbolizes the operating part, the physical support of the function while the upper one is the driver of this physical part. In the middle, an icon symbolizes the interaction between these two parts. A number in the lower right corner identifies the box in the diagram and its location in the hierarchy. Its form is 0XXX.Y where 0XXX represents the successive indices of the "mother boxes" and the Y the index of the box in the current diagram.

- Arrows, whose aspect is significative of the flow they represent : bold for physical flows, normal for information flows, and dashed for control flows. These arrows are used either to link "boxes" or to materialize the importation of an existing flow from an upper level. In this case, they are linked to a plain circle materializing the link with the hierarchy.

To improve the description accuracy, textual templates have been added to each graphic element. These templates do not have to be filled completely in the first steps of the design process, this has to be done only when useful.

1.2 *Using this method*

The use of this method is quite similar to the use of functional ones. First, the designer has to build a context diagram. As for SADT, this one locates the system in the outside environment. To do so, it shows the function performed by the system and all the flows implied in the problem (see example on fig 2).

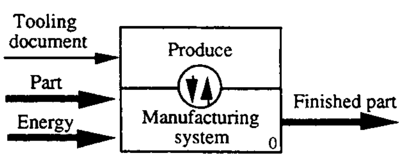


Figure 2: Three post model, diagram 0

The next step consists in dividing the zero diagram in order to refine the description. The examples below (figures 3 to 4) show the first description levels of a small tooling model used as a laboratory validation.

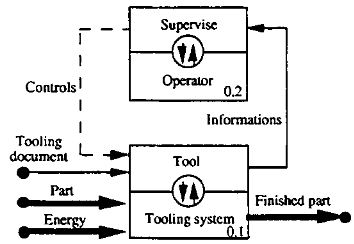


Figure 3: Decomposition of diagram 0

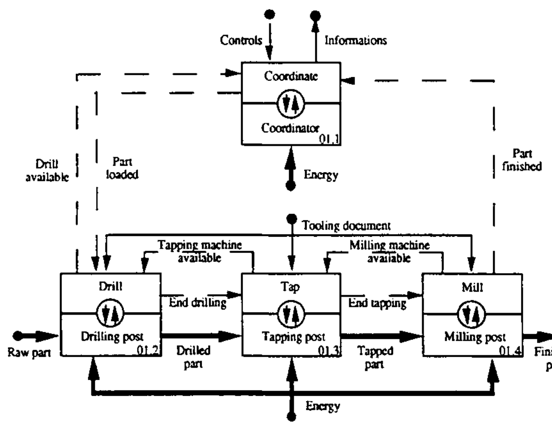


Figure 4: Decomposition of diagram 0.1

This description can be done down to the lower abstraction levels of the system allowing the representation of the physical parts implementing it.

2. VALIDATION APPROACH

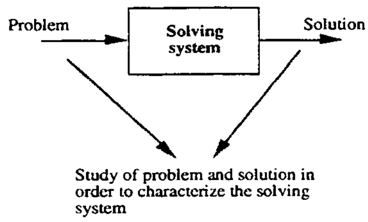
2.1 *Aims*

In order to define the experimental context, we searched the main topics that were relevant to the usability and interest of an AMS design method. In our context, two main topics can be found:

- *Capacity of the design system* (designer + method). This topic will help to know if the use of a particular method allows the designer to increase his capacity of design problem solving and, if so, in what extent.

- *Interaction between the method and the designer.* This topic will help to find out whether the design helping tool is well suited to designers working habits.

In front of each topic, it is possible to list several points of interest that characterize it in a more accurate way. As an example, for the topic "capacity of the design system" we can put "solving time", "refinement degree of the solution", ...



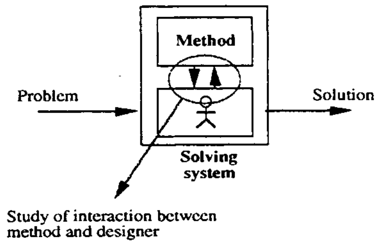


Figure 5: The two main topics retained to characterize a design method.

From these points of interest, it is now possible to build the questions to which the experimentations will have to provide an answer, and then to list the experimental means required to do so.

2.2 *Experimental context*

In order to take into account evaluation criteria and available observation means, we retained the following:

- For objective criteria:
  - the "paper" result of the design activity,
  - a video recording of the designer's activity during the solving of the problem.

- For subjective criterions:
  - an "a posteriori" verbalization stimulated by questions.

Concerning the experimental context, in order to solve the problem they were submitted to, subjects were placed, once at a time, in a calm isolated room. They were given the following equipment:

- the terms of the problem to be solved,
- paper and pens of different colours,
- a small reference document on the method they had to use when it was the case.

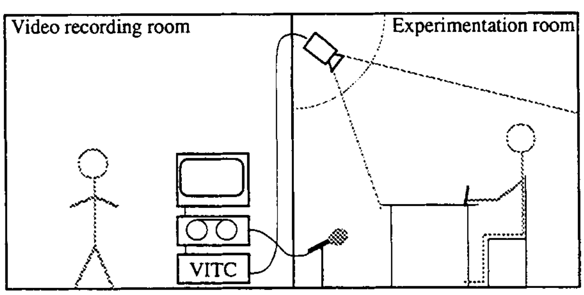


Figure 6: Experimental platform.

Video means used to record designers activity were hidden in the ceiling of the experimentation room and were running without the subjects knowing. This caution has been taken in order to avoid a bias in the experimentations.The subjects could be frightened to know they are being watched (figure 6).

2.3 *Experimental protocol*

In order to reach the aim of the experimentations, which is to study the influence of the use of a design method on the solving process, but also in order to qualify the interest of using a method rather than an other, it is necessary to use three test scenari whose results could be compared. These scenari are detailed below.

The first one deals with the solving of an AMS problem without prescribing the use of a specific method. The subject keeps the free choice of using a method he knows or of not using any method at all. This first test will allow us to assess roughly the experience of each subject in the AMS field and will be, for each subject, the reference point for studying the benefit (or not) of the use of the different methods.

The second scenario deals with the solving of an AMS problem using a well known method: SADT (Structured Analysis and Design Technique). SADT is a commonly used method in this field and its choice for this second test will allow us to quantify the benefit (or loss) resulting in the use of a method regarding the lack of method.

The third scenario deals with the solving of an AMS problem using the AMSD method. This third test aims at assessing the interest of using AMSD in this context (has its use had a positive influence on the problem solving ?), but aims also at revealing possible defaults AMSD could present in order to propose improvements.

The carrying out of these scenari has required a training period from the sujects to the two prescribed methods. After this period, all subjects selected for this experimentation (6 subjects) were supposed to have an homogeneous knowledge on the methods (SADT & AMSD).

Then, each one of the six subjects realized three tests, using the three scenari successively to solve three different problems. To minimize inter-individual differences, problems have been chosen of a same difficulty level and distributed among the three scenari.

After each test, paper results describing the solution to the problems were kept and comments of the subjects collected through questionnaires and a verbalization.

Subjects were also asked to assess the difficulties they felt about the problem using "linear evaluation scales" on the topics: understanding of the problem, reflection, transcript, search in hierarchy levels, check, correction.

Through this experimental protocol and thanks to the experimental platform described above, we extracted the informations that allowed us to deduce results that are introduced in the next part.

3. RESULTS

In this part, we will present a synthesis of the results obtained concerning the proposed solutions and the subjective criterions without details about the characteristics of each method. For a more detailed presentation see Popieul (94).

3.1 Used strategies and solutions proposed by the subjects

First of all, the times needed to reach a solution are rather homogeneous and vary from 1h30 to 2 hours. It must be noticed that the subjects have worked for similar durations but proposed more or less complete solutions. As a matter of fact, subjects that didn't feel any difficulties to propose a solution went to a further level of detail than subjects who were in a failure or semi-failure situation which gave-up after a same amount of time.

The satisfaction of the prescribed constraints was found rather independent of the method and specific to each subject. Mainly, all strong constraints have been taken into account (definition of the system, sequence of actions, ...). Light constraints have been more or less taken into account (system low cost, system easyness, ...). The explanations of the subjects on this point lay in the difficulty encountered during solving. If a solution which has been found with difficulty meets the strong constraints, this solution is kept, even regardless the light constraints, subjects preferring a less suited solution than a failure situation (cognitive cost (Wisser 88)).

Concerning proposed solutions, it must be noticed that AMSD does not increase the quality of the proposed system in itself (the idea). This seems rather logical and is confirmed by the fact that the quality of solutions is homogeneous for each subject and does not evolve with the used method. Most often, the best solutions are proposed by subjects having a good experience of industry, which confirms the great part of re-use in design.

However, what is true for the global system is false for the sole driving part of the system. As a matter of fact, AMSD improved the quality of the driving part of the proposed system. Only AMSD incited the subjects in sharing the driving of the system among sub-systems. With SADT or without any method, driving has always been considered as centralized and working on a set of inputs and outputs concerning the whole system.

We also noticed that the use of AMSD helped the description quality of the proposed solution. Indeed, the number of understandable solutions that could be easily implemented is higher with AMSD (5/6) than with SADT (3/6) or without method (3/6). This is mainly due to the subjects sudden awareness of the necessity to describe the operative part both with the control part and not only one or the other, which is mainly what they did in the two other scenari.

3.2 Influence of the methods during the design phase

Subjective evaluations with questionnaires. The most interesting point concerning these evaluations is that three out of the six subjects found that the problem they solved with AMSD was the easiest among the three they had to solve, which was wrong because they were dealing with two different problems (problems 1 & 2).

As this feeling cannot be explained by a real difference in problem difficulties (these ones being almost the same, except for the third problem, which was perceived as more difficult) it can probably be explained by a higher easiness of ideas transcription when using AMSD than in the two other cases.

Except for little problems when dealing with conveyors, the first results about the use of AMSD are positive. First of all, following the satisfaction of use expressed in the assessment questionnaires, but also, in a more objective way, as a consequence of the way the method has been used. Most of the subjects (5/6) worked with AMSD in the same manner:

- construction of a draft representation of the proposed system,
- realization of an AMSD model of this draft,
- refinement of this AMSD model (each of the boxes) leading to the identification of missing informations that are reintroduced in the upper levels (addition of sensors, flows, ...).

Though we thought they would proceed as follow:

- description of the problem with AMSD,
- refinement,
- design of subparts (physical ones) of the proposed system in order to implement functions of the identified modules,
- modelling of these solutions with AMSD,
- refinement of these solutions and description of the corresponding driving part.

Subjects justified their way of proceeding by the easiness of the problems they had to solve. However, they also noticed that for a larger scale problem, they would also make very soon a draft representation of the physical solution.

Subjective evaluation of difficulty levels. Before starting a global study of all the topics used to characterize the different points of difficulty perceived by the subjects during the experimentations, first, we focused on both topics revealed by the analysis of the questionnaires: reflection and transcription. For each of these two topics, we built a contingency table summing up the 18 tests made with the three methods. To do so, the difficulty perceived by the subjects has been equally divided into three ranges : low, medium and high. Resulting tables are introduced below (table 1).

Table 1: Contingency tables for reflection and transcription topics

Reflection	Low	Medium	High
SD	1	4	1
SADT		5	1
AMSD	3	3	

Transcription	Low	Medium	High
SD	1	1	4
SADT	1	2	3
AMSD	4	2	

These tables put forward that globally, reflection and mainly transcription difficulties have been perceived lower when using AMSD than in the two other cases. This confirms the results deduced from the answers to questionnaires.

In order to assess globally this difference, we computed the inter-subjects means of the subjectives evaluations for each topic for the three methods. The resulting graph is presented below (figure 7).

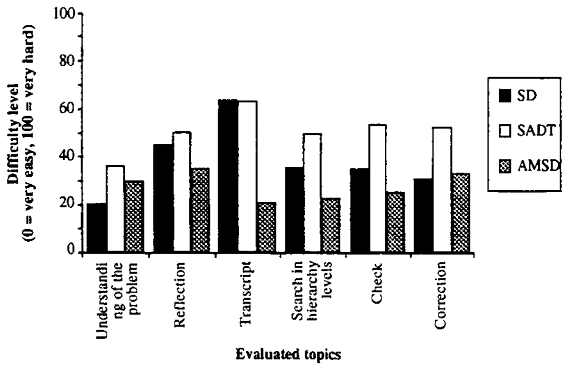


Figure 7: Inter-subjects means of subjectives evaluations

It appears clearly on this graph that, on average, AMSD reveals to be more efficient than other methods concerning transcription and does not increase the difficulty linked to reflection.

Concerning the other evaluated topics, we can notice that SADT gets systematically worse results. This confirms the low fitting of this method to the context of AMS design, which has been already shown in a previous work (Popieul 94).

The medium results obtained by AMSD on these other topics are the same that those obtained by a work without any prescribed method and even a little better.

3.3 *Activity analysis*

Raw data. From a preview of the video tapes recorded during each test, we extracted the few basic activities that can be observed as consequences of the mainly cognitive task induced by design problem solving. Nine activities have been found :

- reading of the statements of the problem,
- reflection,
- problem representation,
- reading of the method user's manual (in case of a scenario with a prescribed method),
- transcript (of final document),
- draft
- reading/checking (draft or final document),
- correction,
- other.

Then each of the 18 video tapes has been reviewed in order to draw the activities chronograms of each test (see an example on figure 8).

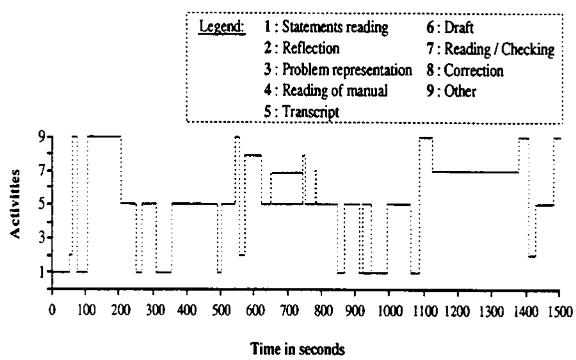


Figure 8: Example of an activity chronogram.

It must be noticed that the attribution of a basic activity at a given time is rather difficult, it is even impossible sometimes to distinguish an activity among several ones. In these cases (rather few in number), several simultaneous activities have been mentioned, supposing an homogeneous time sharing between them. This happened mainly for :

- reflection & transcript,
- reading/checking & transcript,
- statements reading & reflection.

Results. From these data characterizing the sequencing of the identified basic activities, several studies have been done. We will present here only the results concerning the ratios of elementary tasks durations regarding the whole test duration for each subject and each test (see example on figure 9).

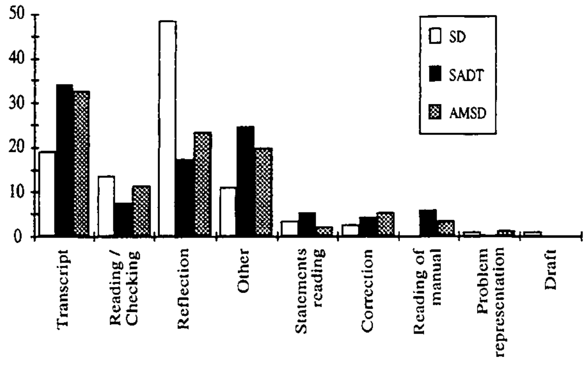


Figure 9: Percentages of total time spent for each basic activity (subject 5, scenari 1, 2, 3).

The first remark before any interpretation is about the difficulty level of the three problems. As problems one and two have been perceived of an equivalent difficulty level and leads to similar reflection ratios, problem three, which has been felt to be more difficult leads to longer reflection durations and was certainly so.

Taking into account this remark, it appears that the time ratio dedicated to reflection is almost the same for the three methods. From this, we can deduce that, as AMSD does not induce a time ratio for reflection higher than the other two scenari, it has no negative influence on time ratio.

The ratio of transcript is the next characteristic point. We can see here that although that AMSD has the heaviest graphic representation, the amount of time dedicated to transcript is similar during its use as for SADT. Even more if we notice that subjects who spent more time for transcription with AMSD are the ones that used more often the given reference manual. This can reveal a little lack of practice with the method.

It can also be noticed that none of the scenari had a real influence on time ratios for reading, checking or correction.

As a synthesis, this temporal study can be concluded in saying that without increasing amounts of time nedeed for reflection or transcript, AMSD allows to reach a higher number of usable solutions.

5. CONCLUSION

The aim of this paper was to present a new design method for Automated Manufacturing Systems and its evaluation from the user point of view.

The way this evaluation was made allowed us to put forward several assets and also several needs of improvements of the method we propose, which reveals the interest of proceeding this way.

However, the experimental context we used remains far from a real AMS design context and thus has to be improved. Our current work is in this way.

At mean terms, simultaneously with the development of AMSD (extension of the existing software support tool with a simulation layer), we intend to realize experiments with a team of designers working on a real AMS problem.

REFERENCES

Lhoste P., Panetto H., Morel G., Colombari G. (1991). Une méthode de spécification et de conception de la partie commande des machines et systèmes automatisés de production manufacturière. *Conf. Int. Génie Automatique et Production Manufacturière,* 12-13 Mars.

Popieul J-C. (1994). Contribution à l'étude et à l'implémentation d'une démarche de conception de machines et systèmes automatisés de production. *Thèse de doctorat,* Valenciennes, le 1er Février.

Smith P.G. (1988). Winning the product race rate. *Machine Design,* May 12.

Visser V. (1988). Giving up a hierarchical plan in a design activity. *Rapport de recherche INRIA N°814,* Mars.

Zaytoon J., Niel E., Mille A., Jutard A. (1993). A temporal SADT for automated manufacturing systems. *International Conference on IEPM,* Mons, Belgium, 2-4 June.

FROM FIELD–BASED STUDIES TO MODELS TO DECISION AIDS – AN APPROACH FOR SUPPORTING HUMAN DECISION–MAKING IN ADVANCED MANUFACTURING SYSTEMS

S. Narayanan

Department of Biomedical & Human Factors Engineering
Wright State University
Dayton, OH 45435
snarayan@valhalla.cs.wright.edu

T. Govindaraj, Leon F. McGinnis, & Christine M. Mitchell

School of Industrial & Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332–0205

Abstract: This paper presents a methodology for supporting human decision–making in advanced manufacturing systems. The methodology is comprised of three components: 1) a conceptual framework to organize descriptions of human role in a manufacturing shop floor, (2) a computational architecture to represent both the automated aspects of a discrete–parts manufacturing system as well as the associated operator problem solving, and (3) guidelines for translating model results into operational decision aids. The methodology is illustrated with an application to electronic assembly systems.

Keywords: Manufacturing systems, automation, computer simulation, object modelling techniques, man/machine interaction, information technology.

1. INTRODUCTION

Modern manufacturing systems are characterized by increasing levels of automation and integration, decreasing levels of work–in–process inventory, and a supervisory role for humans (Cohen and Zysman, 1988; Jaikumar, 1986). Modeling and analysis of the system to gain a better understanding of system complexities are critical for design and ongoing system management. Although the applications of the various modeling methods, such as analytical models and discrete–event simulation models, have provided useful insights to specific questions about the system, such methods have largely ignored the role of people. As a result, the role of the human in supervisory control of advanced manufacturing processes is uncertain.

A growing body of recent literature in *user–centered design* has begun to study the role of people in predominantly automated manufacturing systems. Such research efforts include social science studies which provide insights on problems with poorly automated systems (Corbett, 1988; Zuboff, 1988), conceptual design–oriented research which provide domain–independent guidelines on systems design (Rouse, 1988), laboratory–based studies that attempt to provide answers to specific questions related to the role of the human operators in manufacturing systems (Benson et al., 1992; Dunkler et al., 1988), and field–based studies that attempt to provide insights on problems based on empirical data in operational facilities (Bereiter and Miller, 1989; Cohen et al., 1992). Currently, however, there appears to be a lack of a unified framework to integrate these results or to facilitate their transfer into operational computer–based aids to support human decisions in advanced manufacturing systems.

This paper presents an approach for supporting human decision–making in advanced manufacturing systems. Based on case studies in real–world manufacturing plants, this research seeks to answer several questions including the following: (1) what are the tasks performed by the shop floor human operators in advanced manufacturing systems? (2) how to develop models of the system and operator functions in a computational platform? and (3) how to translate model results to better design the automated aspects of the system and to develop off–line and on–line computer–based decision aids for supporting decision making.

This research has three major components: (1) a conceptual framework to organize descriptions of human role in a manufacturing shop floor, (2) a computational architecture to represent both the automated aspects of a discrete–parts manufacturing system as well as the associated operator problem solving, and (3) a set of guidelines for translating model results into operational decision aids. These three components are illustrated with their application to electronic assembly systems.

2. THE CONTEXT: ELECTRONIC ASSEMBLY SYSTEMS

Figure 1 depicts a typical printed circuit card assembly line. Unpopulated printed circuit card boards (PCBs) are manually placed on the screen printer where solder paste is applied on them. The boards then go to the placement area. One or more automated placement machines place various components including resistors, capacitors, and integrated circuits (ICs) on the board. The boards then pass through a reflow oven and reach the insertion area. One or more automated insertion machines insert through–hole components on the board. Large through–hole components which could not be inserted by the automated insertion machines are inserted manually. The boards then go through a wave solder process. The boards are then cleaned, sheared, and manually touched up, if necessary. Testing and repair comprise the final stages of the process. Boards must pass both the in–circuit test (ICT) and functional test before being sent for packaging.

The above description is for a generic electronic assembly line. There may be minor variations to the system depending on the type of boards produced in the manufacturing plant. For example, an assembly system producing boards with only the surface mount components will not have any automated insertion machines.

Although people in the administrative, technical, and managerial divisions of a company all play an integral role in the successful operation of an electronic assembly enterprise, this research focuses on shop floor personnel that include technicians, engineers, and first–line managers in the factory. Supporting human role in these systems involves specification of the human functions in the system operation and control. The methodology outlined in this paper supports understanding and aiding of human role in manufacturing systems.

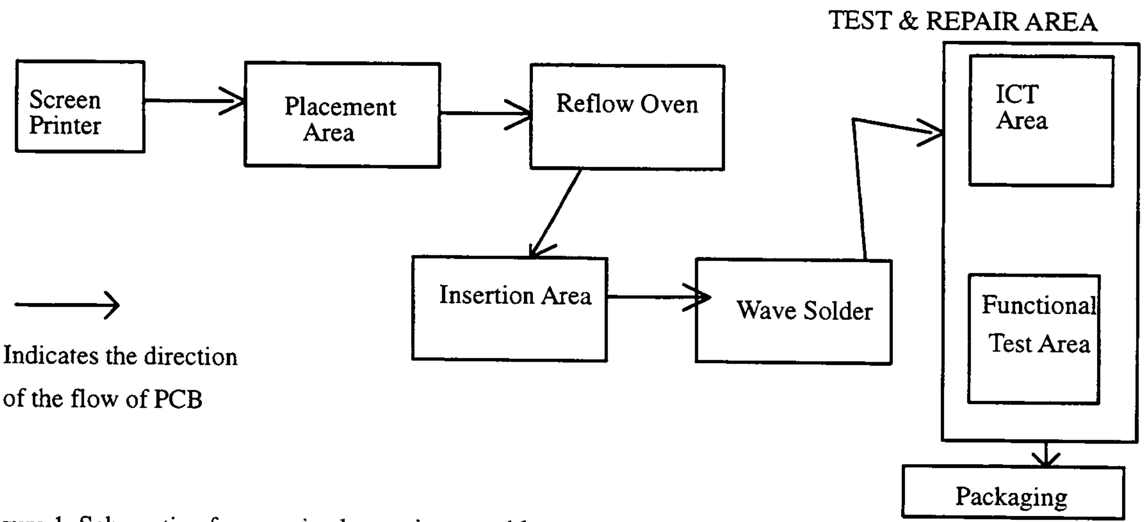


Figure 1. Schematic of a generic electronic assembly system.

3. THE METHODOLOGY

The methodology comprises of three phases: (1) performing field study analysis, (2) developing computer models, and (3) implementing decision aids. The key issues in each of the three phases and the support mechanisms are outlined in Table 1.

An important outcome of performing field study analysis is the development of an organized description of the activities performed by operators and automated components of a system. The field study phase is supported through a conceptual framework that helps generate a system description in terms of the manufacturing processes, operations performed at the various stages of the process, and tasks performed by people and automated equipment on the shop floor.

The fundamental issue in modeling is one of representation. Abstractions are provided in a computational architecture designed to support modeling of relevant aspects of automated activities and operator problem solving in a system. The architecture is implemented in an object–oriented environment; it facilitates modeling and analysis of a system (Narayanan et al., 1994).

Table 1. Components of the methodology.

Phases	Key Issues	Support Mechanism
Field Study Analysis	Organized description of operator functions and automated activities	Conceptual framework
Modeling	Representation of automated activities and operator problem solving	Computational architecture
Decision Aids	Identification of task demands and approaches to support	Implementation guidelines

In designing and developing decision aids, the key issue is to identify task demands based on the results of modeling. This phase is supported through implementation guidelines. The three support mechanisms are illustrated by application to electronic assembly systems.

3.1 Conceptual framework

An operator/automation function decomposition of the various stages of electronic assembly systems is shown in Table 2. In the table, the columns represent the type of operators or automated equipment, and the rows represent the various stages of the manufacturing process. TL denotes technicians with low skill level and training, TH denotes technicians with high skill level and training, ET denotes technically functioning engineers, EM denotes managerially functioning engineers, or first line managers, and AUTO denotes automated equipment.

The various stages of the manufacturing process correspond to the schematic of Figure 1. All the sections between the major areas shown in the schematic is grouped together as transit in Table 2.

The descriptive framework highlights the manual and automated tasks performed in the system, necessary skill levels, and potential support areas. For example, it can be inferred from the table that technicians with low skill levels and training apply solder and load the PCB in the solder area, monitor the reflow oven and the wave solder process, move material in the transit areas, and log in data about the PCB. Tasks performed by the first line managers are system wide and include shift supervison, demand forecasting, manpower scheduling, and job release batching. Automated equipment in the form of a conveyor moves material through the assembly line. Vision systems and automated testers function in the test and repair area. Automated placement and insertion machines populate the PCB with components.

The second phase in the proposed methodology involves the development of computer models of relevant aspects of the system. Modeling is supported by a computational architecture which is described in the following section.

3.2 Computational architecture

The computational architecture designed to support this methodology has domain–specific constructs for modeling manufacturing automation and task–specific abstractions for representing operator problem–solving.

Table 2. An operator/automation function decomposition in electronic assembly systems.

Type	TL	TH	ET	EM	AUTO
Screen Printer	Apply solder paste Load board			Job Release	
Placement		Monitor	Process planning for sequencing and feeder component allocation (most)	Resource Allocation, Man power scheduling, layout design	Component placement Automated process plan generator (some)
Reflow Oven	Monitor				Move, Heat
Insertion		Monitor Component insertion	Process planning	Demand forecasting	Component insertion
Wave solder	Monitor				Move, Process
ICT	Data log	Troubleshoot, Repair			Tester, Vision
Functional Test		Troubleshoot, Repair	Troubleshoot		Test support
Transit	Move material, clean, shear				

The architecture was applied to three major modeling problems in electronic assembly systems: (1) analysis of concurrent placement machines, (2) analysis of a conveyor–based assembly line, and (3) analysis of troubleshooting in an assembly line. The three applications are outlined in this section.

Concurrent placement machines: The goal of this case study was to develop a descriptive model of machines that emulate various characteristics in real–world placement or insertion machines used in printed circuit–card assembly. The time taken by a machine to place all the required components on a board is the *cycle time* for that assembly. Process planning, performed by a human operator or by an automated program, attempts to minimize the cycle time. Decisions made in process planning that affect the cycle time include *arragement* and *sequencing* decisions that address the arrangement of component feeders and sequencing of the placement operations (McGinnis, et al., 1992).

Analytic models attempt to generate optimal solutions for the component feeder assignment and the placement sequence. The optimum cycle time estimates obtained by these algorithms may or may not be the actual cycle time one encounters when the solution is applied to real problems. The solution depends on the assumptions made in the model and the approximations made in the solution procedure. Models which emulate the characteristics of a placement machine are necessary to validate solutions obtained from analytical models. The architecture was used to develop generic abstractions to model a broad class of such machines. Models developed using the architecture can also be used for what–if analysis in evaluating between various possible solutions. Details of the implementation of the architecture for this application can be found in Narayanan et al. (1994).

Conveyor–based assembly line: This case study focused on the modeling and analysis of automated material flow control. The system modeled was a conveyor–based electronic assembly line used in a high volume automotive application. In contrast to the placement machine model described earlier, this case study focused on a higher–level shop floor control. In this application, conveyor control plays an important role in the performance of the overall system. The architecture provided controller objects which enabled direct mapping to controllers in the actual system and facilitated the testing of different control strategies on the system performance (Sreekanth, 1994). First–line managers in an electronic assembly line can potentially use the architecture to model and analyze similar systems and make decisions on resource allocation, batching, and layout design.

Electronic troubleshooting: This case study involved the modeling and analysis of the performance of operator troubleshooting in the in–circuit test area of an assembly plant. Cohen et al. (1992) describe an operator function model of electronic troubleshooting based on real–world data collection. That study and subsequent research on problem solving and learning in this task provided some interesting insights (Narayanan et al., 1992; Ram et al., 1994). First, operators rely significantly on past experiences; for example, they recollected known problems while troubleshooting. Second, operators relied predominantly on *shallow reasoning methods* using heuristic and context–sensitive associative knowledge during problem solving. Third, the observation by Barr and Feigenbaum (1981), that humans often solve a problem by finding a way to think about a problem that facilitates a solution, was also evident. Finally, learning plays an important role in the increased performance of an operator with experience.

Although troubleshooting operators are trained for performing skill–oriented actions such as replacing a missing part through manual soldering, there is little on–line or off–line computer–based support for problem diagnosis in extant electronic assembly systems, clearly a critical phase in troubleshooting. The model developed for this application contains abstractions for representing various types of associative and heuristic knowledge and learning strategies used by operators in troubleshooting.

3.3 Implementation Guidelines

The guidelines to implementing decision aids are specific to the types of users and the functions they perform in electronic assembly systems. The guidelines outlined below are tailored to the three applications discussed in the previous section.

The modeling results of the placement machine analysis are perhaps most useful to a process planners. A front–end support system could potentially enhance the utility of the model for what–if analysis. The support system should ease the burden of system specification for the process planner, support visualization of the model results, and highlight the concurrency of the interactions between the components of the machine being modeled. The current implementation of the architecture outputs Gantt–chart like graphs on the state of the various components of a machine over time.

The modeling results of the conveyor–based assembly line are most useful to first–line managers. Aggregate systems–level analysis can be performed and decisions on system redesign and resource allocation can be made based on the model results. A graphical configuration support system could be useful in facilitating system specification. In the current implementation of the architecture, a Motif–based graphical front–end supports system configuration.

The modeling results of research on electronic troubleshooting and learning has implications on developing both off–line computer–based tutoring systems and on–line computer–based decision support systems. Major issues in developing an intelligent tutoring system include what to teach and how to teach (Psotka et al., 1988). Modeling results indicate that it is valuable to teach shallow troubleshooting knowledge, including context–specific associative knowledge and general heuristic knowledge. Similarly, one means of providing on–line decision support is by having a repository of past cases, appropriately indexed so context–sensitive assistance can be provided during problem solving.

4. SUMMARY

Understanding the role of automation and people is critical in designing advanced manufacturing systems. In the context of electronic assembly systems, a methodology was presented to organize descriptions of the role of shop floor personnel and automation in a system, to develop computer models of relevant aspects of the system, and to translate model results into operational decision aids. The constructs in the computational architecture to support operator problem solving are limited to material flow control and troubleshooting tasks. Further research is necessary to develop abstractions for planning and design. Future research will also investigate extensibility of the methodology to other advanced manufacturing systems.

REFERENCES

Barr, A., and E. A. Feigenbaum (Eds.) (1981). *The Handbook of Artificial Intelligence*, **1**, Addison–Wesley, Reading, MA.

Benson, C. R., T. Govindaraj, C. M. Mitchell, and S. P. Krosner (1992). Effectiveness of direct manipulation interaction in the supervisory control of FMS part movement. *Information and Decision Technologies*, **18**, pp. 33 – 53.

Bereiter, S. R., and S. M. Miller (1989). A field–based study of troubleshooter's information utilization in computer–controlled manufacturing systems. *IEEE Transactions on Systems, Man, and Cybernetics*. **19**, no. 1, pp. 205 – 219.

Cohen, S. M., C. M. Mitchell, and T. Govindaraj (1992). Analysis and aiding the human operator in electronics assembly. In *Design for Manufacturability: A System Approach to Concurrent Engineering and Ergonomics.* (M. Helander and M. Nagamachi, (Eds.)). London: Taylor & Francis, pp. 361–376.

Cohen, S. S., and J. Zysman (1988). Manufacturing innovation and American industrial competitiveness. *Science.* **239**, pp. 1110 – 1115.

Corbett, J. M. (1988). Human centered advanced manufacturing systems: From rhetoric to reality. *International Journal of Industrial Ergonomics.* **5**, pp. 83 – 90.

Dunkler, O., C. M. Mitchell, T. Govindaraj, and J. C. Ammons (1988). The effectiveness of supervisory control strategies in flexible manufacturing systems. *IEEE Transactions on Systems, Man, and Cybernetics.* **18**, no. 2, pp. 223 – 237, March/April.

Jaikumar, R. (1986). Post industrial manufacturing. *Harvard Business Review.* pp. 69 – 76, November/ December.

McGinnis, L. F., J. C. Ammons, M. Carlyle, L. Cranmer, G. W. DePuy, K. P. Ellis, C. A. Tovey, and H. Xu (1992). Automated process planning for printed circuit card assembly. *IIE Transactions*, **24** (4), pp. 18–30.

Narayanan, S., D. A. Bodner, U. Sreekanth, T. Govindaraj, L. F. McGinnis, & C. M. Mitchell (1994). An object–based, direct–image approach to the modeling and simulation of manufacturing systems. Technical Report, *MHRC–TR–94–04,* Material Handling Research Center, Georgia Institute of Technology, Atlanta, September.

Narayanan, S., A. Ram, S. M. Cohen, C. M. Mitchell, and T. Govindaraj (1992). Knowledge–based diagnostic problem solving and learning in the test area of electronics assembly manufacturing. *Proceedings of the SPIE's Applications of Artificial Intelligence X Conference.*

Psotka, J., L. D. Massey, S. A. Mutter (Eds.) (1988). *Intelligent Tutoring Systems: Lessons Learned,* Lawrence Erlbaum Associates, Inc., Hillsdale, New Jersey.

Ram, A., S. Narayanan, and M. T. Cox (1994). Learning to troubleshoot: Multistrategy learning of diagnostic knowledge for a real–world problem–solving task. To Appear in *Cognitive Science.*

Rouse, W. B. (1988). Intelligent decision support for advanced manufacturing systems. *Manufacturing Review,* **1**, no. 4, December.

Sreekanth, U. (1994). POEMS: A process and object environment for manufacturing simulation. Unpublished *master's thesis*, Georgia Institute of Technology, School of Industrial and Systems Engineering, Atlanta, GA.

Zuboff, S. (1988). *In the Age of the Smart Machine.* Basic Books, Inc., New York.

Modeling Human Performance of Intermittent Contact Tasks

J. Won* **and N. Hogan****

**Massachusetts Institute of Technology, Department of Mechanical Engineering, Cambridge, MA, USA. E-mail: juiceman@mit.edu*

***Massachusetts Institute of Technology, Departments of Mechanical Engineering and Brain and Cognitive Sciences, Cambridge, MA USA.*

Abstract. The work presented here tests the validity of a model of human performance of intermittent contact tasks in which the arm is alternately free and constrained by contact with an object. Experiments were performed to quantify typical human behavior while making unconstrained, constrained, and intermittently constrained two-joint arm movements. Additional experiments were performed which measured the static two-joint impedance of the experimental subjects. A series of linear and non-linear time-invariant models were created from these impedance measurements and then tested for competence in modeling human behavior measured in the three tasks. It was found that tasks with simple dynamics such as the unconstrained and constrained movements could be competently modeled with the proposed models. However, none of the models tested could match the behavior of the arm during the intermittent contact task.

Key Words. Arm Movements, Modeling, Impedance, Human-Machine Interface, Simulation.

1. INTRODUCTION

Dynamic models of human manipulation are important for the design and control of machines that interact with humans, such as haptic virtual environment displays, telerobots and artificial limbs. Intermittent contact is commonplace in these tasks but has not yet been the subject of extensive study. Due to the abrupt change of mechanical conditions between constrained and unconstrained motion, they present a challenging test of theories of how the brain controls the arm and hand. Previous work by Flash (1987) showed that a simple time-invariant model can be remarkably competent in reproducing unconstrained human arm movement. The focus of this work is to determine if that time-invariant model (or a simple extension of it) could also reproduce the behavior of the arm when it is constrained by contact with an object.

The models tested are based on the postulate that movement generation and control of contact are both generalizations of posture. This so-called "equilibrium-point" theory of human motor control is based on the observation that muscles coupled with their reflex loops behave like viscoelastic springs. Acting in groups (including agonist and antagonist) they define an equilibrium or attractor posture of the limb. Neural commands that define a sequence of postures along a desired trajectory (a virtual trajectory) will generate spring-like muscle forces to track the moving attractor point. To control contact and manipulation, moving the attractor point within an object will cause the limb to exert a force on that object. The appealing feature of this theory is that dynamics have been mapped onto statics; movement and manipulation may be represented and executed in terms of postures without the complexities of multi-body spatial linkage dynamics.

In prior work, Flash showed that a simulation model based on this theory was remarkably competent to reproduce unconstrained human arm movement. In those simulations, Flash modeled the limb as a nonlinear inertial linkage and used a linear model with stiffness and damping for the net impedance of the human arm, but varied its shape and orientation in proportion to the measured variation of arm stiffness with posture. Using straight line virtual trajectories as the driving input, Flash found that the simulations were competent in recreating even the subtle details of observed human arm trajectories.

We extend Flash's work to kinematically constrained motions of the arm. Data was collected as human subjects made reaching movements while attached to a simple planar robot. On selected trials a constraint was applied by the robot immobilizing one of its degrees of freedom. In a further experiment, subjects making reaching movements were initially constrained by the robot, but the constraint was removed in mid-movement. This task was akin to a human making intermittent contact with a hard surface during object manipulation.

Taking this data, a series of simulations were performed in the hopes of finding a single model which could encompass all three cases. Three different models were tested - a linear impedance model similar to the one used by Flash, a nonlinear stiffness model derived from measurements of postural impedance, and a nonlinear stiffness and damping model.

2. MATERIALS AND METHODS

Data from two separate experiment procedures are used in this paper. The first experiment measures the postural stiffness field of the subject's arm. The second experiment measures the behavior of the arm as it performs simple point-to-point reaching movements. Perturbations are applied to the movements in the form of positional constraints in order to simulate full and intermittent contact.

The postural impedance measurement is similar to the technique used by Mussa-Ivaldi et al. (1987) and Shadmehr et al. (1993). The movement data was taken from earlier work presented in Won (1993). A short summary of the data collection technique follows. A more detailed explanation of the movement experiments can be found in Won (1993).

2.1. *Experimental Procedure*

Subjects were required to perform both experiments in the same day. First the postural measurement were taken, and then after a half hour rest the movement experiments were performed. In both experiments, subjects were seated in front of a planar two-link robot grasping the force transducer handle attached to the end-effector. The arm of the subject was suspended with a sling so that the upper arm and forearm remained in plane with the arms of the robot. In addition, the subject wore a wrist cuff which effectively immobilized the wrist leaving only the shoulder and elbow as available degrees of freedom. Position of and force at the handle was recorded during both experiments.

Postural Measurement. After grasping the handle of the robot, the subject was instructed to position the handle under a target and close their eyes. The robot then applied a series of 35 displacements in different directions from this initial spot. The displacements were achieved by commanding the robot to gradually change its desired position in the form of a ramp taking 9 seconds to achieve the final displacement 5 cm away. After each displacement, the motors were turned off and the arm was allowed to return to the initial position before the next displacement was applied. The subject was instructed to ignore the perturbation and "not to intervene voluntarily."

Movement Experiment. For the movement data, the subject started with the handle under an illuminated target light was prompted to make a movement when the light at the current position was extinguished and another light at the target position was illuminated. The subject was instructed to move to the target as if reaching for some object at that point. The subject was also instructed to attempt to ignore any abnormalities during the movement and perform the movements normally.

At intervals, perturbations were applied to the subject through one of the magnetic brakes. These perturbations took two forms. The first type locked the brake associated with the elbow joint of the robot from the start of the movement to the end. Locking the elbow removed a degree of freedom from the robot so that the handle path was constrained to lie on a circle. The constraint was applied in such a way that both the start and target positions lay upon the circular path. The second type of constraint locked the elbow joint at the beginning of the movement. However, during movement, the constraint was released without warning, and the arm was free to move in both degrees of freedom.

2.2. *Model of the Human Arm*

The model of the human arm has two major components. The first is a dynamic model of the inertias of both the human arm and the robot arm. Assuming that the distance from the robot base to the shoulder of the subject is fixed, this is a two degree of freedom system. The second component is the input, $\boldsymbol{\tau}$, the vector of torques applied to each joint of the human arm by the muscles. By the virtual trajectory hypothesis, it will be assumed that the behavior of the human produced torques can be described by the following relationship.

$$\boldsymbol{\tau} = \boldsymbol{\tau}_k(\boldsymbol{\theta} - \boldsymbol{\theta}_o) + \boldsymbol{\tau}_b(\dot{\boldsymbol{\theta}} - \dot{\boldsymbol{\theta}}_o) \qquad (1)$$

where the central command of the CNS is represented as a motion of an attractor point $\boldsymbol{\theta}_o$. Previously the static and viscous relationships, $\boldsymbol{\tau}_k$ and $\boldsymbol{\tau}_b$, have been modeled as a linear stiffness and damping.

In this paper, linear and nonlinear models based on field approximation were explored. In the experiment described previously, force and position data were recorded as each subject was displaced from posture. This data describes a vector field of forces as a function of x-y displacement as shown in Figure 1(a). Since Mussa-Ivaldi showed that the stiffness tensor was more invariant in joint coordinates than in endpoint coordinates, the endpoint vector field was transformed to joint space as shown in Figure 1(b) before attempting to model the field.

Using the technique developed by Mussa-Ivaldi et al. (1991), the force fields were approximated by a linear combination of "basis fields" centered at different locations of the data workspace. Previous work has shown that the shoulder-elbow postural field is well approximated by a conservative field (Mussa-Ivaldi et al., 1991; Shadmehr et al., 1993). Therefore the "basis field" used is the gradient of the bivariate Gaussian radial basis function suggested by Poggio and Girosi (1990).

$$G_i = \exp\left(-\frac{(x-x_i)^2}{2\sigma_x^2} - \frac{(y-y_i)^2}{2\sigma_x^2}\right) \quad (2)$$

The basis field is then

$$\phi_i = \nabla G_i \quad (3)$$

so that the force at a given displacement is approximated by the n basis fields.

$$\mathbf{F} \approx \sum_{i=1}^{n} c_i \phi_i \quad (4)$$

The set of coefficients, c_i, are found by a least squares fit by using the pseudo-inverse to find the solution to the following equation.

$$\mathbf{f} = \mathbf{\Phi c} \quad (5)$$

where $\mathbf{f}$ is the vector of the measured forces and $\mathbf{\Phi}$ is the matrix of basis fields evaluated each data position.

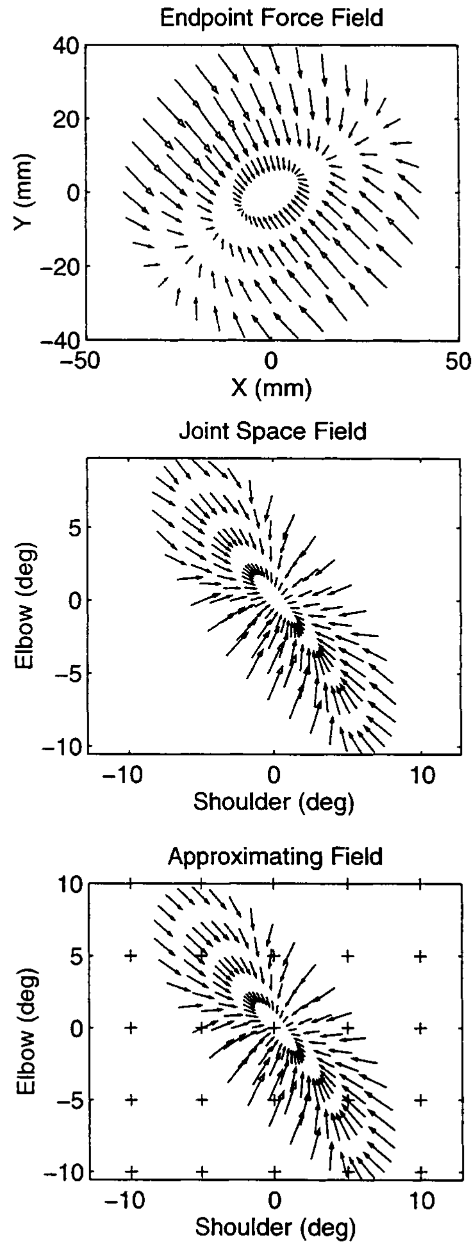


Fig. 1. Static force field measurement. (a) The experimental data as the hand is steadily displaced over a 5 cm range. The vectors represent the forces measured at the displacement marked by the tail of the vector. (b) The field in (a) transformed to shoulder-elbow space. The vectors now represent torque magnitude and direction. (c) The analytical model of (b) using 25 basis fields centered at the locations marked by a cross.

2.3. *Simulations*

A 4th order Runge-Kutta numerical integrator was used to integrate the equations of motion for this system. The muscular torques were generated by moving the equilibrium point in Equation 1 along a straight line from the start to the target with a minimum jerk velocity profile (Flash and Hogan, 1985).

In addition, the Broydon-Fletcher-Goldfarb-Shanno method of constrained non-linear optimization was used to tune the simulations to match the measured data. It was shown in the work of Mussa-Ivaldi that coactivation causes only

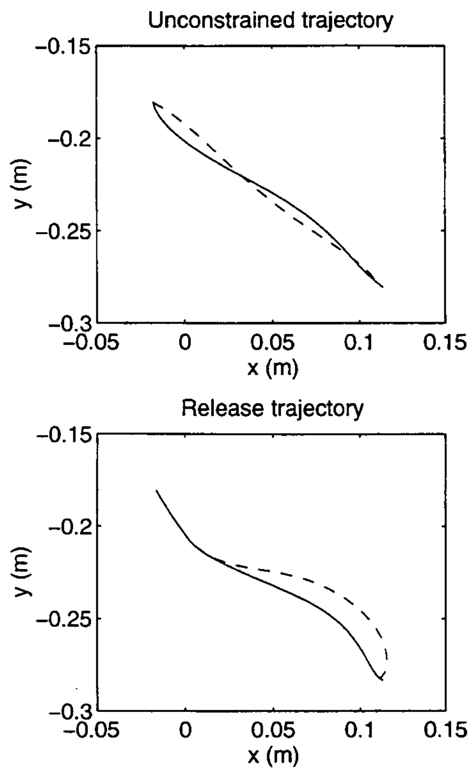


Fig. 2. The result of the linear impedance simulations for the (a) free and (b) intermittent contact cases. The dotted line represents average experimental data while the solid line is the simulation output. Movements start from the upper left proceeding diagonally downward.

a scaling of stiffness, and although not experimentally verified, it will also be assumed that only a simple scaling occurs with the viscous component as well. Therefore, the optimization tuned the scaling factor of the static and viscous impedance components of the arm to minimize a squared error between the simulated trajectory and the measured trajectory. In order to enforce agreement across all three cases, the cost function was actually the sum of the errors of the two kinematically free cases - the unconstrained and the constrained and released case.

3. RESULTS

3.1. *Linear Simulations*

Using force data for displacements within 1 cm of posture, endpoint stiffness was found by linear regression and transformed (using the Jacobian) into joint space. Due to the lack of reliable measurements of the two-joint viscous component, the linear viscous impedance was assumed to have the same shape as the linear stiffness but with a different scaling factor. For the subject whose data will be shown in this paper, the total linear impedance is

$$\tau = \alpha \begin{bmatrix} -24.1 & -8.5 \\ -8.5 & -36.0 \end{bmatrix} \begin{bmatrix} \theta^s - \theta^s_o \\ \theta^e - \theta^e_o \end{bmatrix} + \beta \begin{bmatrix} -24.1 & -8.5 \\ -8.5 & -36.0 \end{bmatrix} \begin{bmatrix} \dot\theta^s - \dot\theta^s_o \\ \dot\theta^e - \dot\theta^e_o \end{bmatrix} \quad (6)$$

where α and β are the scaling factors tuned by the optimization routine.

When simulating unconstrained movements, this linear impedance was very successful in recreating the experimental data (see Flash (1987) and Won (1993)). However when attempting to fit all three experimental cases, the limits of this model are made clear. Figure 3(b) displays forces that are much larger than the typical forces seen in the experimental data of Figure 3(a). The maximum force measured experimentally was under 4 Newtons instead of the 6 Newtons predicted by the simulation. In addition due to the need to minimize errors in the trajectories shown in Figure 2, the optimization reduced the damping to a level such that the unconstrained trajectory in (a) is less straight than the experimental data and the release trajectory in (b) has a smaller inward return. This was a common problem as the system required larger damping for straight unconstrained trajectories but smaller damping for the large inward returns. The parameters themselves are physically justifiable since the scales, $\alpha = 1.89$ and $\beta = .07$, give a time constant of 37 ms which is near the 50 ms twitch contraction time of human biceps and triceps (Buchthal and Schmalbruch, 1970).

3.2. *Nonlinear Stiffness*

Taking the measured joint space fields and using 25 evenly-spaced basis fields, an analytic approximation to the field was calculated as shown in Figure 1(c). Similar to the linear simulations, this field will be scaled by a constant α by the optimization routine. In contrast to the linear approximation, the field approximation includes important nonlinearities such as the drop in stiffness at larger displacement as well as the nonsymmetrical behavior of the actual field. The viscous component was linear as in the previous case.

Using the non-linear field, the performance of the simulation in the fully constrained case (Figure 3(c)) was much improved. Due to the non-linear stiffness, the forces evoked by the constraint are similar in magnitude (and direction) to the experimental data. However, as in the linear case, the attempt to match both the unconstrained case and the release case failed (Figure 4). Again, the 39 ms time constant ($\alpha = 2.07$ and $\beta = .08$) is close to a realistic value for human muscles.

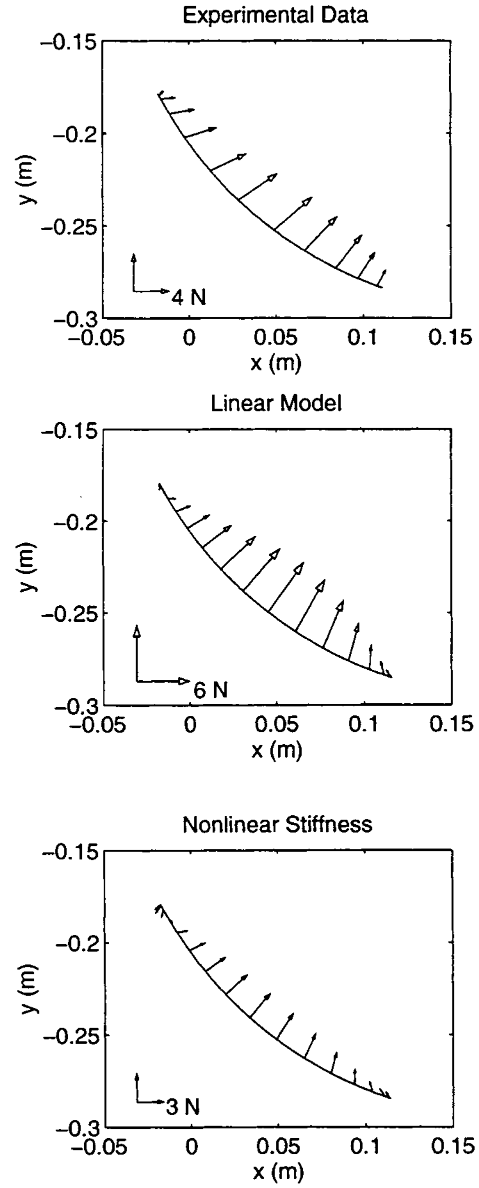


Fig. 3. The constrained motion data for the linear and nonlinear simulations. The constraint forces the hand to move along a circular arc. The measured forces are indicated by vectors. (a) is typical data from a human subject performing the task. (b) is the result of the linear impedance simulation. (c) is the result of the nonlinear stiffness, linear damping simulation. Movements start on the upper left.

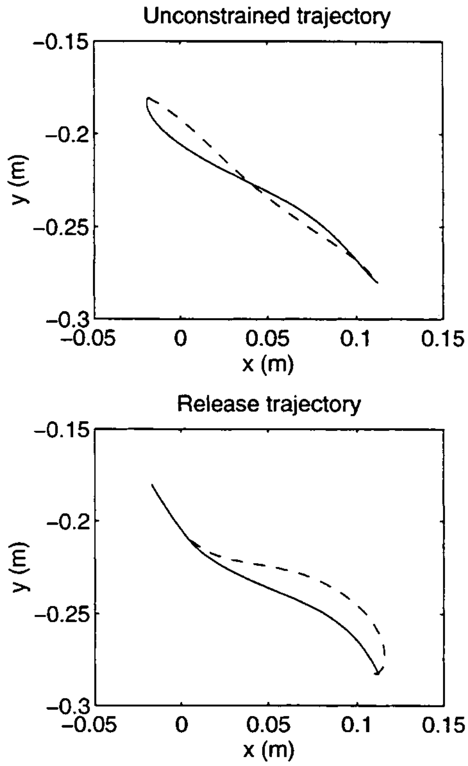


Fig. 4. The result of the nonlinear stiffness simulations for the (a) free and (b) intermittent contact cases. The dotted line represents average experimental data while the solid line is the simulation output.

3.3. *Nonlinear Damping*

These results suggested that the different cases required different damping values. Although encouraging straighter unconstrained trajectories, higher damping hinders the arm from making large inward returns after release from the constraint. We next attempted to use a nonlinear damping model. Unfortunately, this impedance component has been virtually unexplored especially in the two-joint case. We made the plausible (but unsupported) assumption that the nonlinearity of the force-velocity relation was similar in form to that of the force-displacement relation. The final model used the approximation for stiffness and the same for damping. With this model, the damping factor will drop at higher velocities which will permit a larger inward return after release when the system is moving at large velocities.

This model did exhibit an inward return comparable to experimental observations but introduced other problems as shown in Figure 5. As the attractor point reaches the target the system is moving quickly. Due to the nonlinearity the viscous field is providing very little resistive torque in this case and permits exaggerated and unrealistic oscillations at the end of the movement. α in this

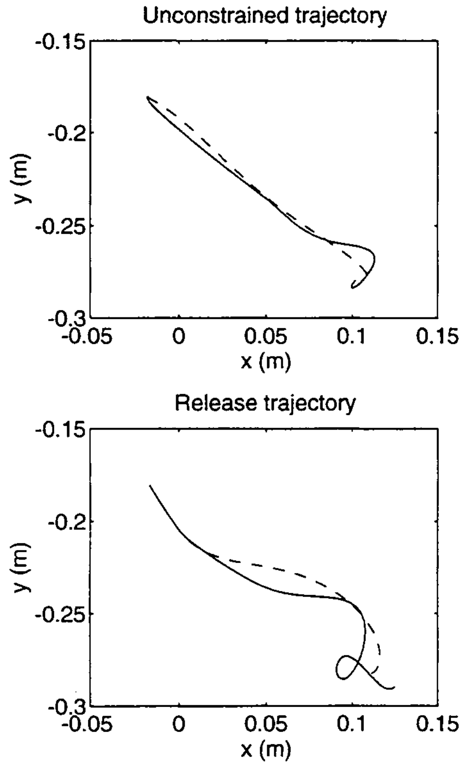


Fig. 5. The result of the nonlinear stiffness and damping simulations for the (a) free and (b) intermittent contact cases. The dotted line represents average experimental data while the solid line is the simulation output.

case was 3.8, and β was 0.8.

4. CONCLUSION

This work extended Flash's simulations of free motion to include sustained and intermittent contact. Using a non-linear model of the postural field greatly improved the accuracy of the sustained contact simulations. However, none of the models investigated provided a satisfactory simulation of the intermittent contact task, primarily due to a lack of reliable data on the subject's viscous behavior. Given the remarkable success of these models in free and sustained contact movements, we recommend that this analysis be continued with a more complete measurement and parameterization of human arm impedance.

ACKNOWLEDGMENTS

This work was supported in part by NIH grants AR-40029 and NS-09343. J. Won was supported by a Department of Defense NDSEG fellowship.

5. REFERENCES

Buchthal F. and Schmalbruch H. (1970) *Contraction Times and Fiber Types in Intact Human Muscle.* Acta Physiol. Scand. 79:435-452.

Flash, T. (1987) *The control of Hand Equilibrium Trajectories in Multi-Joint Arm Movements.* Biological Cybernetics 57:257-274.

Flash, T. and Hogan N. (1985) *The Coordination of Arm Movements: An Experimentally Confirmed Mathematical Model.* J. Neuroscience 5(7):1688-1703.

Mussa-Ivaldi F.A., Hogan N., and Bizzi E. (1985) *Neural, Mechanical, and Geometric Factors Subserving Arm Posture.* J. Neuroscience 5:2732-2743.

Mussa Ivaldi F.A., Bizzi E., Giszter S.F. (1991) *Transforming Plans into Actions by Tuning Passive Behavior: A Field Approximation Approach.* In: Proceedings of the IEEE International Symposium on Intelligent Control pp. 101-109.

Poggio T. and Girosi F. (1990) *A Theory of Networks for Learning.* Science 247:978-982.

Shadmehr R., Mussa-Ivaldi F.A., and Bizzi E. (1993) *Postural Force Fields of the Human Arm and Their Role in Generating Multi-Joint Movements.* J. Neuroscience 13(1):45-62.

Won J. (1993) *The Control of Constrained and Partially Constrained Arm Movements.* MIT MS Thesis.

ANALYSIS OF THE HUMAN OPERATOR CONTROLLING A TELEOPERATED MICROSURGICAL ROBOT

Lynette A. Jones and Ian W. Hunter,

Department of Mechanical Engineering,
Massachusetts Institute of Technology,
Cambridge, MA 02139 USA

Abstract: In the design and construction of force–reflecting interfaces the properties of the human operator using the interface must be considered. Two components of the human operator system are of particular importance, namely the biomechanical and sensory processing subsystems. In order to feed back information about contact forces and the mechanical properties of objects being manipulated by the robot it is essential to know the operator's sensitivity to changes in force, stiffness and viscosity and the gain and bandwidth required to feed back this information. Psychophysical studies have demonstrated that the ability of human operators to resolve changes in stiffness and viscosity is inferior to that for force. This means that if an operator controlling a robot is required to detect changes in stiffness or viscosity, then small changes will have to be scaled up in order for the operator to detect them consistently.

Keywords: Human–machine interface Human perception Operators Robotics Telemanipulation Teleoperation

1. INTRODUCTION

Information about the performance of a human operator controlling a robot can be displayed in a variety of ways, although typically it is presented visually. Considerable effort has therefore been devoted to the development of head–mounted displays and other visual devices, such as high–resolution rear projection screens, that provide the operator with a realistic image of the working environment of the robot. With the development of teleoperated robots that are used to perform high precision tasks, such as microsurgery, it has become evident that visual feedback alone is inadequate, and that task performance can be enhanced if the operator has information about the forces experienced by the robot. Force–reflecting interfaces have therefore been incorporated into a number of preliminary surgical robotic systems as they have been shown to be important to the detection of contact conditions with body tissues and to the control of cutting forces (Hunter et al., 1993; Kazanzides et al., 1992; Sabatini et al., 1989; Shimoga and Khosla, 1994).

In order for robotic systems to be accepted in surgery, they must at the very least provide surgeons with the capabilities that they already have, that is, information about contact between the surgical instrument and the tissue being manipulated, the forces generated at the interface between the tool and tissue (which can be in the order of 0.3–0.4 N), and the position of the surgical instruments (Charles et al., 1989). Because the forces generated by robots performing high precision tasks are typically very small, they must be scaled up before they are transmitted back through a computer to a force–reflecting interface in order for the operator to perceive them. This means that a teleoperated microsurgical robot can not only enhance the performance of surgeons by extending their dextrous capabilities but also provide them with information that is usually absent during an operation.

The design of force–reflecting and haptic interfaces used to control and sense the status of a teleoperated robot must take into account the biomechanical and sensory properties of the operator. In order to feed back information about contact forces and the mechanical properties of objects being manipulated by the robot it is essential to know the operator's sensitivity to changes in force, stiffness and viscosity and the gain and bandwidth required to feed back this information. These data can only come from psychophysical studies of the human proprioceptive and tactile systems.

2. HUMAN OPERATOR SYSTEM

The human operator who forms part of the closed–loop system controlling the robot is usually considered as a lumped system rather than as a summation of a number of subsystems, some of which are shown in Figure 1. Two of these component subsystems are particularly important to the design of force–reflecting interfaces used in microsurgical robots, namely the limb mechanics and sensory processing units. Biomechanical studies of the human arm and forearm are essential because of the large changes in both the force and stiffness of the hand and arm during high precision tasks. By increasing the stiffness of the hand and arm the position resolution of the hand improves, which makes it less susceptible to the effects of a perturbing force (Tendick and Stark, 1989).

As the forces used to grasp a surgical tool increase, so too does the stiffness of the forearm and hand (Jones and Hunter, 1990a), which means that the perceived magnitude of forces reflected back from a robot to the operator will change as a function of the stiffness of the operator's arm. Measurements of forearm stiffness indicate that during high precision tasks involving relatively small finger forces, this increase can be in the order of a 2–3 times that of the stiffness of the limb at rest. These variations would result in large differences in the amplitude of forces experienced by an operator through a force–reflecting interface. If these variations in the mechanical properties of the arm are not incorporated into the design of the interface, an operator would be unable to make accurate judgments of the forces experienced by the microsurgical robot and under some conditions would be unable to perceive the forces.

Analysis of the sensory processing subsystem of the human operator has become an integral part of the development of the microsurgical robot. A number of psychophysical studies have been conducted in which the sensitivity of subjects to variations in mechanical variables has been measured. The objective of these experiments is to determine the consistency with which an operator can detect changes in stiffness and viscosity. Anecdotal reports suggest that the latter two variables are used by surgeons when resecting tissue and that it may therefore optimize operator control if information about these variables is presented to the surgeon controlling the microsurgical robot.

Two methods have been used to determine the sensitivity of subjects to changes in stiffness and viscosity. The first is the contralateral limb matching procedure, and the second is the adaptive psychophysical technique known as the transformed up–down method (Gescheider, 1985).

3. METHODS

3.1 Apparatus

Psychophysical studies have been performed using a custom–built experimental rig (1 m square at the base and 2 m high) on which two electromagnetic linear motors are mounted, each of which is powered by a custom 2000 W water–cooled class A/B low noise power amplifier. Connected to the translation stage of each motor is a rod and aluminum cuff that is used to clamp the motor to the wrist. Mounted in the center of the rig is a four–axis electrically adjustable seat, and in front of this chair is a horizontal beam on which aluminum elbow supports are bolted.

Force is measured using a load cell (Entran Devices) mounted at the end of the brass tube to which the wrist is attached. An inductive displacement transducer (Data Instruments) is affixed behind the linear motor coil coaxially with the tube attaching to the wrist. Displacement, force and motor current signals are sampled at 5 kHz by 18–bit A/Ds (Hewlett–Packard) resident on a VXI bus in an I/O crate controlled by an IBM R6000/320 computer. The motor stiffness and viscosity is controlled digitally via 16–bit D/As (Tasco) running at 5 kHz.

3.2 Procedure

Subjects sit in the chair with each elbow joint resting in an individually molded plastic cast (Aquaplast) that is in turn encased in an aluminum elbow support. The position of these supports is adjusted until the angle between the upper arm and forearm is 90°. During the experiments subjects wear headphones

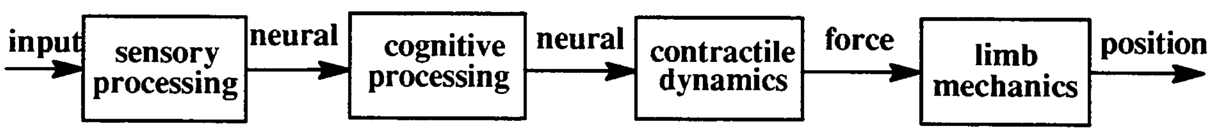


Fig. 1. Human operator subsystems

through which tones are presented that signal the start and termination of each trial. To prevent subjects from using visual cues they are blindfolded during the experiments.

Four separate experiments have been conducted to measure the sensory resolution or differential thresholds for stiffness and viscosity. In two experiments the contralateral matching procedure was used and results from these have been published (Jones and Hunter, 1990b, 1993). In two new experiments reported here the transformed up–down method has been used to study the perception of stiffness and viscosity. The former method required that subjects adjust the stiffness or viscosity of a linear motor connected to one arm, the matching limb, until it was perceived to be the same as that of the motor attached to the other, reference, arm. The matching stiffness or viscosity was controlled by rotating the left foot (about the ankle axis) on a bi–directional footplate attached to the rig, on which an angular position transducer was mounted. The stiffness or viscosity of the motor attached to the reference arm was fixed on each trial at one of 10 values and ranged from 0 to 6260 N/m in the stiffness experiment and from 2 to 1024 N.s/m for viscosity. Each of these amplitudes was presented ten times. With this procedure the subject has active control over the mechanical variable of interest and can adjust it continuously until the stiffness or viscosity of the two motors is perceived to be the same.

The transformed up–down procedure differs from the matching technique in that subjects do not control the stiffness or viscosity of the motors but indicate on each trial which motor has the larger stiffness or viscosity. A single–interval forced–choice paradigm was used in these experiments. Subjects were given 8 s to move their arms against the motors and then indicate (by pulling the appropriate arm towards them) which motor had the greater stiffness or viscosity. The correct response was indicated to the subject by a pulse delivered to the correct arm. For each block of 50 trials the stiffness or viscosity of the reference signal was fixed at one of 6 or 8 values, respectively. For stiffness this ranged from 200 to 6400 N/m and from 4 to 512 N.s/m for viscosity.

At the start of each block of trials the comparison signal was set to twice that of the reference signal and thereafter the difference between the reference and comparison signal decreased or increased depending on the subject's responses. A transformed up–down procedure was used (Levitt, 1971) to track the subject's threshold, with the absolute difference between the reference and comparison signals being halved after two successive correct responses and doubled after one incorrect response. These rules seek a stimulus level that corresponds to 71% correct performance. Because it was the difference between the reference and comparison signals that was relevant, it was randomly determined on each trail whether the difference was added to or subtracted from the reference to create the comparison signal. In order to control for errors of measurement that can arise when stimuli are presented to different receptive areas on the body, such as the left and right arms, the reference and comparison signals were randomly assigned to the left and right motors on each trial.

4. RESULTS

In the matching experiments the mean and standard deviation of the ten matching stiffnesses and viscosities at each reference amplitude were calculated for each of the ten subjects. Differential thresholds can be calculated from matching data provided that the stimulus is continuously variable and that both the standard and comparison stimuli can be presented simultaneously (Gescheider, 1985). The threshold is defined as the standard deviation of the matching stiffness or viscosity values produced at each reference viscosity, and the ratio of the differential threshold to the matching stimulus is known as the Weber fraction. This fraction is an index of sensory resolution and can be compared across different variables. The Weber fractions for stiffness and viscosity estimated using the matching technique are shown in Figures 2 and 3. As can be seen in the figures, the Weber fraction increases considerably at lower stimulus levels, a feature often noted in other sensory systems. Over the range of stimulus intensities that the Weber fractions were essentially constant, the mean value for stiffness was 0.23 and for viscosity it was 0.34.

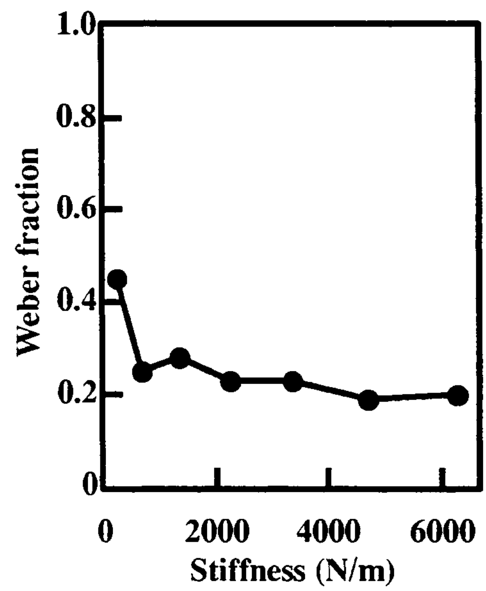


Fig. 2. Weber fraction (SD/mean matching stiffness) measured at each matching stiffness. Each point represents the mean of 10 subjects' data.

The transformed up–down method tracks the differential threshold on the basis of the subject's responses and is therefore considered to be a more efficient method of measuring thresholds. In general the absolute difference between the reference and comparison signals decreased during the first 10–15 trials and thereafter hovered around the subject's threshold level. The differential threshold was defined as the geometric mean of the absolute difference between the reference and comparison signal for the last 30 trials and was calculated for each subject at each reference amplitude. The relation between the reference stiffness and viscosity and the differential threshold are shown in Figures 4 and 5.

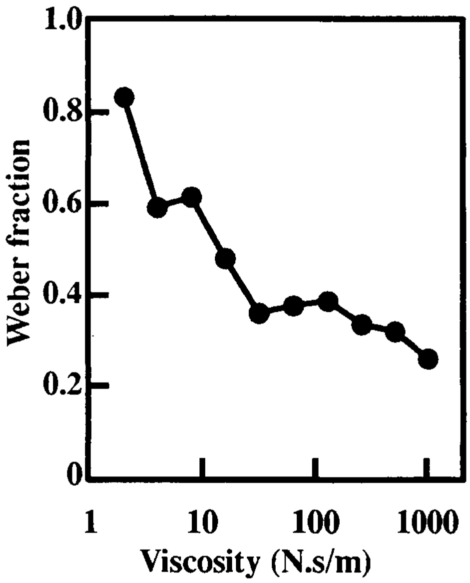


Fig. 3. Weber fraction (SD/mean matching viscosity) measured at each matching viscosity. Each point represents the mean of 11 subjects' data.

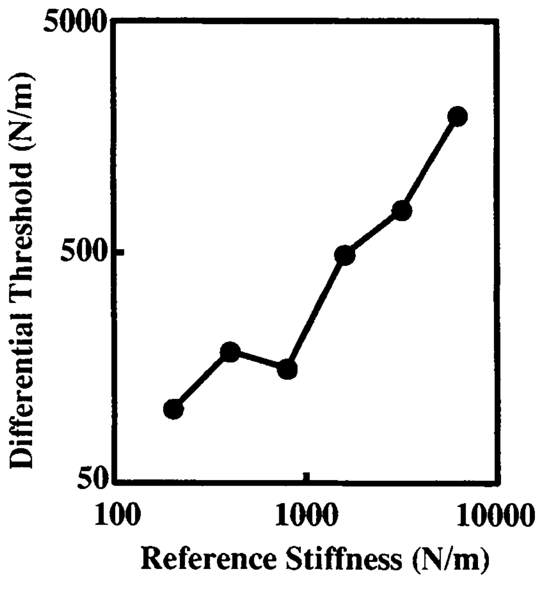


Fig. 4. Differential thresholds measured at each reference stiffness. The slope of the line is the Weber fraction for stiffness.

As can be seen in Figures 4 and 5 the thresholds increase monotonically and essentially linearly with the amplitude of the reference stimulus. The linear regression line fitted to these data accounted for 99% of the variance in Figure 4 and 98% of the variance in Figure 5. The slope of this line is the Weber fraction, and for stiffness it was 0.16 and for viscosity 0.19. These data suggest that there is a considerable improvement in the ability to detect changes in stiffness or viscosity if the task is simply to indicate which of two mechanical systems has the larger stiffness or viscosity and that if the element of control is involved in perception, as is the case with the matching technique, then performance deteriorates.

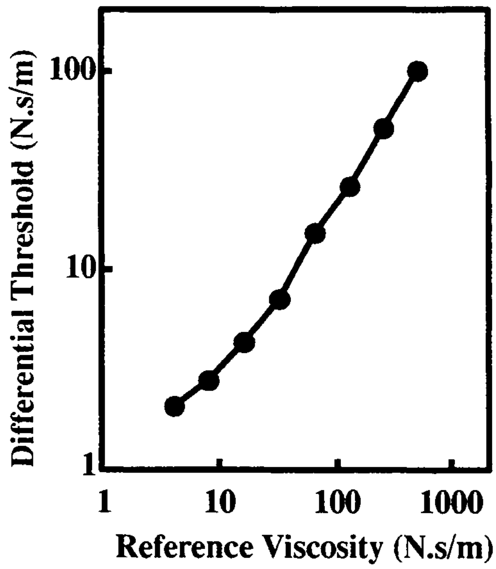


Fig. 5. Differential thresholds measured at each reference viscosity.

5. CONCLUSIONS

The Weber fraction represents the magnitude of change in a variable that an operator can reliably discriminate. The larger the fraction, the greater the amplitude change required for discrimination. The results presented here indicate that subjects are better at discriminating changes in stiffness and viscosity than has been reported previously using other psychophysical methods (Jones and Hunter, 1990b, 1993), but that for both of these variables relatively large variations must occur before an operator will detect them. The ability to discriminate these changes is, however, a function of the current level of stiffness or viscosity being fed back from the environment, as Figures 4 and 5 show. This means that if an operator controlling a robot is required to detect a change in stiffness or viscosity (e.g. moving through an environment of varying viscosity or detecting a change in the consistency of tissue that may be indicative of disease), then small changes would have to be scaled up in order for the operator to detect them consistently.

The ability to discriminate changes in stiffness and viscosity is considerably inferior to that reported for force. Estimates of the Weber fraction for force range from 0.05 to 0.10 (Jones, 1986; Tan et al., 1992). This means that at a relatively low force level of 500 mN a change of 50 mN or more can be perceived.

The studies described here are relevant to the design of haptic and force-reflecting interfaces in that they provide information about how sensitive operators are to changes in the mechanical properties of objects in the environment with which either they or a robot are interacting. In addition, they indicate how this information can be presented to operators in order to optimize their performance.

ACKNOWLEDGEMENTS

This research was supported by the Medical Research Council of Canada, the Institute for Robotics and Intelligent Systems (IRIS), a Canadian Network of Centres of Excellence, and the U.S. Office of Naval Research.

REFERENCES

Charles, S., R.E. Williams and B. Hamel (1989). Design of a surgeon-machine interface for teleoperated microsurgery. *Proc. Ann. Internat. Conf. IEEE Eng. Med. Biol. Soc.*, **11**, 883–884.

Gescheider, G.A. (1985). *Psychophysics: method, theory, and application*. Erlbaum, Hillsdale.

Hunter, I.W., T.D. Doukoglou, S.R. Lafontaine, P.G. Charette, L.A. Jones, M.A. Sagar, G.D. Mallinson and P.J. Hunter (1993) A teleoperated microsurgical robot and associated virtual environment for eye surgery. *Presence*, **2**, 265–280.

Jones, L.A. (1986). Perception of force and weight: Theory and research. *Psychol. Bull.*, **100**, 29–42.

Jones, L.A. and I.W. Hunter (1990a). Dynamics of human elbow joint stiffness. *Proc. 16th Can. Med. Biol. Eng. Conf.*, **16**, 175–176.

Jones, L.A. and I.W. Hunter (1990b). A perceptual analysis of stiffness. *Exp. Brain Res., 79*, 150–156.

Jones, L.A. and I.W. Hunter (1993). A perceptual analysis of viscosity. *Exp. Brain Res., 94*, 343–351.

Kazanzides, P., J. Zuhars, B. Mittelstadt and R.H. Taylor (1992). Force sensing and control for a surgical robot. *Proc. IEEE Internat. Conf. Robot. Automat.*, 1, 612–617.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *J. Acous. Soc. Am.*, **49**, 467–477.

Sabatini, A.M., M. Bergamasco and P. Dario (1989). Force feedback-based telemanipulation for robot surgery on soft tissues. *Proc. Ann. Internat. Conf. IEEE Eng. Med. Biol. Soc.*, **11**, 890–891.

Shimoga, K.B. and P.K. Khosla (1994). Touch and force reflection for telepresence surgery. *Proc. Ann. Internat. Conf. IEEE Eng. Med. Biol. Soc.*, **16**, 1049–1050.

Tan, H.Z., X.D. Pang and N.I. Durlach (1992). Manual resolution of length, force and compliance. *Advan, Robot., DSC-Vol* **42**, 13–18.

Tendick, F. and L. Stark (1989). Analysis of the surgeon's grasp for telerobotic surgical manipulation. *Proc. Ann. Internat. Conf. IEEE Eng. Med. Biol. Soc.*, **11**, 914–915.

A MODEL OF THE ARM'S NEUROMUSCULAR SYSTEM FOR MANUAL CONTROL

René van Paassen*

*Delft University of Technology, Faculty of Aerospace Engineering, Kluijverweg 1, 2629 HS Delft
contact person: ir. R.J.A.W. Hosman. Author's present address:
University of Kassel, Human-Machine Systems, Mönchebergstraße 7, 34109 Kassel

Abstract. This paper introduces a model for the neuromuscular system of the pilot's arm. The model is elaborated for roll control of an aircraft with a side stick. Two experiments to test the model and provide parameter values are described. Not all parameters of the model could be individually estimated at the hand of the experimental data. The resulting model and parameter values can however adequately describe manual control with active and passive side sticks, also in cases where external disturbances act on the stick/manipulator system.

Keywords. Manual control, aircraft control, models, manipulators, neuromuscular system

1. INTRODUCTION

The application of side sticks in some new types of aicraft, examples are the General Dynamics F16 and Airbus A320, was an incentive to start research on the manual control with side sticks at the Faculty of Aerospace Engineering, Delft University of Technology. The tool chosen for this work was a servo-controlled side stick. This stick is equipped with a hydraulic servo motor with hydrostatic bearings. In combination with a real-time computer such a stick can be used to simulate the properties, i.e. mass, damping, springs and non-linear dynamics, of a wide range of side-stick hardware.

However, simulation of existing hardware does not exhaust the capabilities of such a servo-controlled side stick. One possibility is the implementation of an active side stick. In this configuration the force exerted on the stick is used as the input of the aircraft, or another system to be controlled, and, using the servo motor, the position of the stick is linked to one of the outputs of the aircraft. A typical example would be roll control input, and feedback of the aircraft's roll rate to the stick roll position (Figure 1). Such an active stick provides the pilot with a wealth of relevant proprioceptive input, which enables him to control the aircraft or simulation with much greater precision (Hosman *et al.*, 1990).

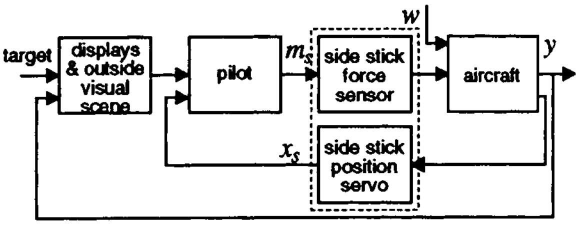


Fig. 1. Control of an aircraft with an active side stick. The moment on the stick, m_s, is used as the aircraft input signal, an output of the aircraft is fed back to the pilot via the stick position x_s. Any disturbance on the aircraft w, e.g. by turbulence, can be felt in the side stick

The moment exerted on the stick (m_s) is the input for the aircraft, an output of the aircraft, e.g. roll rate ($\dot{\phi}$), controls the stick position. One of the challenges a designer of a fly-by-wire flight control system, either with side stick or with a center stick, faces, is the greater freedom to choose the dynamic properties of the manipulator. In order to use this potential, with passive manipulators or with active manipulators as mentioned above, he needs to be able to predict the behaviour of the combination of side stick and the pilot's hand. Data about the neuromuscular system (nms) of the pilot's arm, in a form suited to clarify the puzzles encountered in the research with side sticks at DUT, or to properly explain and predict the behaviour of side sticks currently in use (e.g. roll ratchet in the F16), was not found in the literature.

Therefore, work was started on modelling of the pilot's nms and on the design of experiments to test the model and find the appropriate parameter values. The model and the experiments were elaborated for the case of roll control of an aircraft with a side stick (Paassen, 1994). This model, and

Table 1 List of parameters for the nms hardware.

component	parameter	symbol
CC	damping coefficient	B_m
	max. contractile moment	m_{max}
SEC	spring constant	K_e
PEC	spring constant	K_p
skin	damping coefficient	B_c
	spring constant	K_c
limb	moment of inertia	I_l

two experiments used to derive its parameter values, are described in this paper.

2. MODEL DESCRIPTION

The newly developed model describes the pilot's nms. One part of the model describes the hardware, i.e. muscles, inertia of the arm, skin flexibility. Another part describes the neural circuits whereby the pilot controls his muscles to attain a certain limb position, and the circuits for compensation of disturbing moments on the stick and limb. The fact that the nms and side stick dynamics are not lumped into the same model, and that the reaction to external disturbances on the stick and limb system can be calculated, distinguishes the model from the nms modelling currently in use in the Optimal Control Model (Kleinman *et al.*, 1970) or the cross-over model by McRuer and Jex (1967).

In order to keep the model complexity within bounds, a number of simplifying assumptions is made. The first concerns the manner in which the pilot moves the side stick. It is assumed that in moving the stick he or she keeps the wrist joint stiff, keeps the elbow in a fixed place and rotates the lower arm inward and outward to move the stick. This permits the modelling of the limb's mass with a single moment of inertia.

The muscles linking the shoulder blade, the trunk and the upper arm, and that effectuate the inward - outward rotation of the upper arm and lower arm, are modelled with a single push-pull muscle model[1]. This model is a three-element muscle model, sometimes also called a Hill type muscle model, see Figure 2. This muscle model contains a series elastic component (SEC), a parallel elastic component (PEC) and a contractile component (CC).

[1] Originally, a pair of non-linear models was also considered. However, the data from the experiments did not support the identification of the parameters of this model (Paassen, 1994).

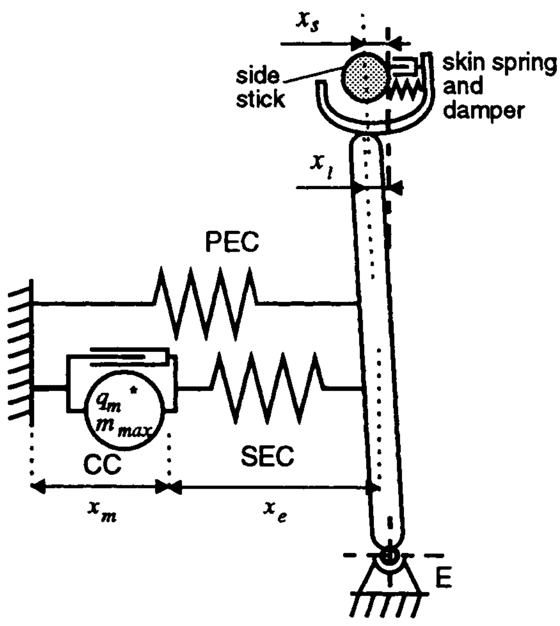


Fig. 2. Schematic top view of the pilot's arm showing the "hardware" of the nms model: a muscle model, limb inertia and a spring and damper modelling skin properties. The figure shows the length of the CC, x_m, extension of the SEC, x_e, limb position x_l and stick position x_s.

The skin is modelled as a flexible link between the limb and the stick consistsing of a spring and a damper mounted in parallel. The skin model is required for obtaining a reasonable simulation when the nms is excited by movement of the stick with the hydraulic actuator, as was done in one of the experiments. A mass, spring and damper analogy for the combination of muscle, skin and limb model is given in Figure 2. Parameters for the skin model, muscle model and the limb inertia are listed in Table 1.

The combination of the limb mass, the muscle and the side stick may be considered as a plant that must be controlled in order to provide the right inputs to the aircraft. In the design of this part of the model is was tried to incorporate the two basic modes of control available to the human. Open loop control, with muscle control signals based on an internal representation (IR) for the execution of fast movements, and closed loop control using sensory information, for execution of precise movements.

A control theory analogy is used to model the control and feedback loops in the nm system. The control signals for the muscle are calculated on the basis of the IR of the limb and the side stick. This internal model is embedded in a closed control loop, see Figure 3. This loop generates both the control signal for the muscle model (q_f) and a reference signal for the expected limb position (x_{l_i}).

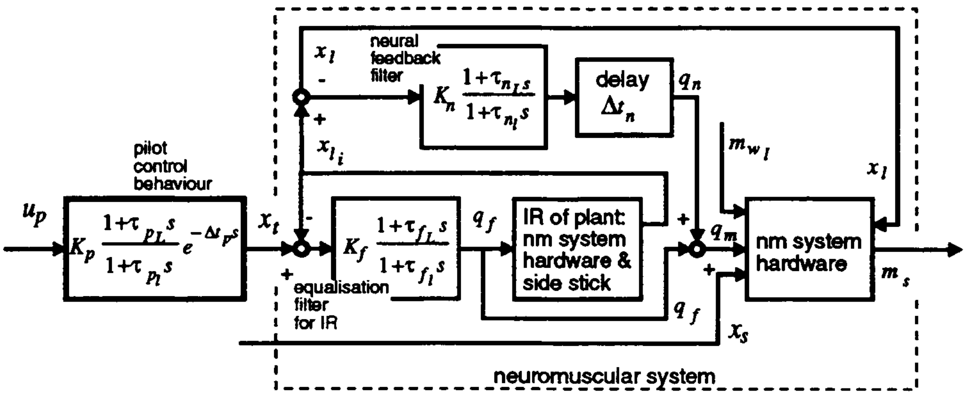


Fig. 3. Combination modelling pilot's control behaviour and the nms. the hardware is described in Figure 2. This figure shows the neural feedback and the equalisation filter for the internal representation (IR). The variable q_m is the control signal for the muscle, built up from the control calculated on basis of the internal model, q_f, and the neural feedback q_n. x_{l_i} is the limb position of the internal model for the hardware, u_p is the visual input of the pilot, m_{wl} is a disturbance on the limb.

The pilot's reactions to disturbances acting on the stick and limb are modelled with a lead-lag filter and a time delay, see also Figure 3 (neural feedback). The time delay models the delay due to transmission times in the axons.

For the complete description of the subject's behaviour in the second experiment, the model has to be combined with a model for the pilot's control behaviour. The model by McRuer and Jex (1967) is used for this purpose. The output of this model is now the target position for the limb, the input is the error perceived on the display.

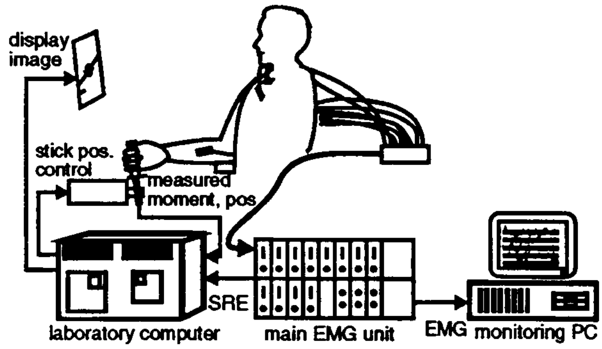


Fig. 4. Set-up used in the experiments.

3. EXPERIMENTS

The model was elaborated for roll control with a side stick. A fixed base set-up, with a hydraulically driven side stick was used to perform experiments that tested the model and provided data for the identification of the parameters (Fig. 4. This paper describes details of two experiments that provided data for the identification of the abovementioned model. A third experiment, calibration of EMG data, is summarily described, for a fourth experiment the reader is referred to (Paassen, 1994).

3.1. *Calibration of the relation between SRE and moment exerted on the stick.*

During all experiments, the subject's EMG signals were measured at seven locations. The EMG measurement unit converted these signals into SRE, smoothed, rectified EMG, by passing the EMG through a full-wave rectifier and a 200 Hz, 3rd order low-pass Butterworth filter. These SRE signals were recorded at a sampling rate of 500 Hz.

For the data processing of the experiments, it was necessary to know the relation between the moment exerted on the stick and the SRE signals. In a calibration experiment, subjects had to press against the stick at 15 force levels, ranging from 30N inward to 25 N outward force on the stick. These levels enclosed the force levels used in the other expeeriments. Results of this experiment permitted interpretation of the SRE signals in the other experiments.

3.2. *Parameter identification of the model components for the muscle, skin and the limb inertia.*

Experiment description. Because the model is quite complicated and contains a sizeable number of parameters, it was attempted to break up the experiments and thus also the parameter identification. In this experiment, only the hardware components of the nms are considered. A test signal, consisting of a pre-defined position displacement, was applied with a hydraulically driven side stick (Figure 5).

Normally, the movement of the stick and the arm would lead to neural reactions, i.e. changes in muscle activation, that try to counteract the movement of the stick. The subjects were asked to suppress these reactions, and thereby allow a

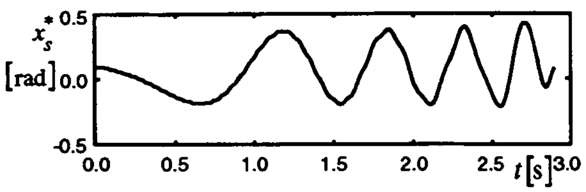


Fig. 5. Input signal for the stick position used in the first experiment

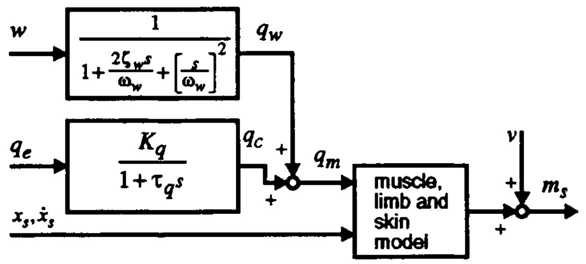


Fig. 6. Model for the estimation of the parameters of the nms hardware. Inputs are the stick position, an input noise w and the activation signal q_e, reconstructed at the hand of the SRE - stick moment calibration. Filter parameters of the noise filter: $\zeta_w = 0.7$ and $\omega_w = 1\,\mathrm{rad\,s^{-1}}$.

measurement of the nms hardware components alone. To help them in this task, feedback of EMG signals was applied. Just prior to a test input, the smoothed rectified EMG signals (SRE) on six muscles were measured (flexors of the wrist, pectoralis major, sternal and clavicular part; extensors of the wrist, deltoid and trapezius). During the test input, the SRE signals were summated in 0.1 s intervals and compared to the levels measured earlier. The difference was squared and summated into an error score, which was presented on a display in front of the subject.

In practice sessions the subjects familiarised themselves with the error score, and learned to suppress a large part of their reactions to the test input, or, more accurately phrased, they learned to keep the activation of the muscles, as measured by the SRE, approximately constant.

When considering only the hardware of the nms, this can be seen as a systems with two inputs, muscle activation and stick position, and one output, the moment exerted on the stick.

For the parameter identification, this two-input model was used. The stick position could be measured without problems. For the second input, the muscle activation, an approximate measure was calculated using the results from the calibration experiment and the measured SRE signals. To express the (large) uncertainty in this activation signal, the model for the identification method also considers a coloured noise at the activation input. The reconstructed activation signal is filtered by

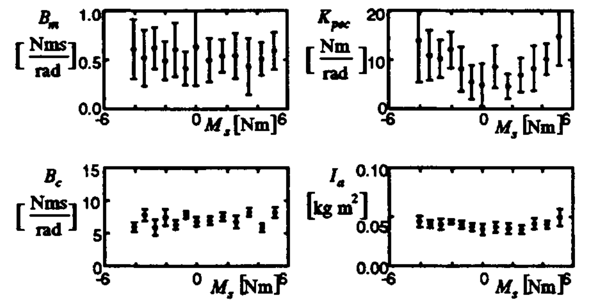


Fig. 7. Results of the parameter estimation of experiment one, four parameters at different offset moments. Male subject, 74 kg, height 1.74 m, lower arm length 0.34 m. The experiment was repeated 10 times, points indicate the average result and bars indicate the sample standard deviation.

a 1st order filter, of which gain and time constant are optimised by the parameter estimation program. In this way the activation signal was used only when this was advantageous for the parameter estimation.

Certain non-linear muscle models, e.g. the model by Winters and Stark (1985), predict a dependency of the (linear) muscle properties on the muscle activation input. To test this, subjects had to exert an offset moment on the stick, prior to and during the test inputs, this was used to control the average activation during a test input.

Results. Figure 7 shows the results of the parameter estimation for four parameters, B_m, K_{pec}, B_c and I_l at different offset moments. The remaining two parameters, K_e and K_c could not be determined in this experiment. The parameter m_{max} acts only as a scale factor, this parameter was set to 40 [Nm]. The SEC spring constant was set to a value $K_e = 50[\mathrm{Nm\,rad^{-1}}]$, roughly based on (Winters and Stark, 1985), and the skin spring constant to, $K_c = 400[\mathrm{Nm\,rad^{-1}}]$. No literature is available for this parameter, but this value assures that the skin compression/extension in the model remains reasonably small. The expected effect of the offset moment on the CC damping coefficient, B_m, was not observed, the offset moment seems only to have a small effect on parameter K_{pec}, something which the non-linear muscle models do not predict; an explanation for this is not known.

3.3. *Measurement of the complete nms & pilot equalisation in a control task.*

In this experiment, the subjects executed a manual control task. They controlled a system that displays the typical characteristics of the aircraft roll dynamics, namely:

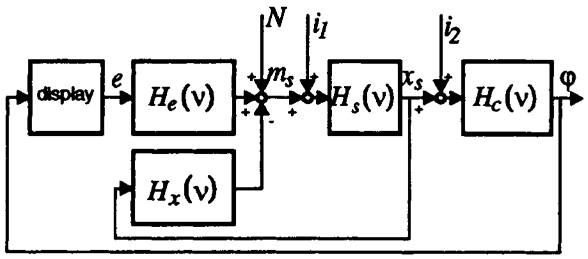


Fig. 8. Control task with two test signals, one (i_1) is a disturbing moment on the side stick, the second (i_2) a disturbance on the controlled system input. This enables the identification of the pilot's moment reponse to stick position, $H_x(\nu)$ and to display presentation, $H_e(\nu)$

$$H_a(s) = \frac{K_a}{s(s\tau_a + 1)} \tag{1}$$

The roll time constant $\tau_a = 0.4$ [s] and the gain constant was $K_a = 4$. In this experiment the pilot, including his nms, is considered to be a two-input, one output system. The inputs are 1) the error presented on the display and 2) the stick position, the output is the moment on the stick. In order to be able to independently identify the reaction to these two inputs, two test signals were used in the experiment (Figure 8). The first signal (i_1) is a disturbing moment on the side stick, the second signal (i_2) is a disturbance on the controlled aircraft-like system. These signals each consist of a sum of fifteen sine functions, chosen so that the subjects cannot recognise any deterministic components in the test signals. The side stick simulation used in the experiment had a mass of 1.5 kg, damping constant 23.3, spring constant 400 N/m. Defined for the centre of the hand, 90 mm above the stick rotation axis.

Take as an example one of the frequencies where test signal i_1 is active. The stick moment, display error and stick position are transformed to obtain the Fourier coefficients M_1, E_1 and X_1 at this frequency. The coefficients M_2, E_2 and X_2 are obtained by interpolation from neighbouring frequencies where the other test signal, i_2, is active. Estimates for the transfer functions for the moment on the stick in response to system error ($\hat{H}_e$) and in response to stick position ($\hat{H}_x$) can be calculated as:

$$\hat{H}_e = \frac{M_1 X_2 - M_2 X_1}{E_1 X_2 - E_2 X_1}, \tag{2}$$

$$\hat{H}_x = \frac{E_1 M_2 - E_2 M_1}{E_1 X_2 - E_2 X_1} \tag{3}$$

It is also possible to calculate an approximation for the standard deviation of these estimates, for this the reader is referred to (Paassen, 1994; Paassen, 1991).

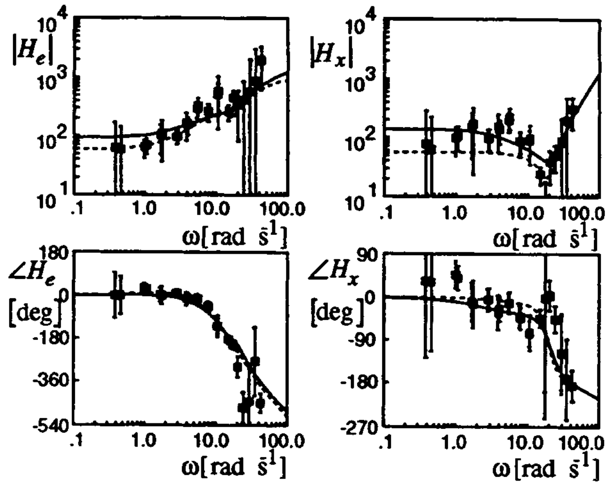


Fig. 9. Results of the second experiment, the estimated describing functions $\hat{H}_e(\nu)$ and $\hat{H}_x(\nu)$, at the frequencies in test signals i_1 and i_2. The vertical bars denote the standard deviation. The dotted lines show the response of the model after parameter fit.

Results. The results of the abovementioned experiment and the subsequent non-parametric estimation are describing functions of the pilot's behaviour, see Figure 9. In a following stage these intermediate results were used to further estimate parameters of the model of the nms. This was done by minimising the weighted, squared distance between the measured transfer functions $\hat{H}_e(\nu)$, $\hat{H}_x(\nu)$ and their model equivalents $\tilde{H}_e(\nu)$ and $\tilde{H}_x(\nu)$ at the fifteen frequencies of i_1 and the fifteen frequencies of i_2. The weighing was done with the inverse of the standard deviations found for $\hat{H}_e(\nu)$ and $\hat{H}_x(\nu)$.

The parameter values that were found for the neuromuscular hardware the previous experiment were borrowed here. This left the parameters from the pilot model (4), the neural feedback model (4) and the neural control model (3) for estimation. The pilot model lead and lag terms were adjusted according to the rules given in (McRuer and Jex, 1967). The other parameters were manually adjusted until a fair match was found between the measured and model frequency responses.

In the parametric estimation of the model with a Gauss-Newton search method, it was not possible to use all parameters in the estimation, see Table 3. The results for the lead and lag terms in the neural feedback are relatively close, suggesting that a proportional feedback and time delay would suffice to model the neural feedback. Lead and/or lag terms for the neural control model could also not be estimated with the Gauss-Newton method.

Table 2 Results from the parameter estimation, first experiment, male subject, 72 kg, 1.75 m, lower arm length, elbow to side stick center 0.31 m.

parameter		value	s.d.
B_m	[Nms rad^{-1}]	0.538	0.241
K_{pec}	[Nm rad^{-1}]	9.14	5.39
B_c	[Nms rad^{-1}]	7.01	1.14
I_a	[kg m^2]	0.0417	0.0059
K_c	[Nm rad^{-1}]	400	(fixed)
K_e	[Nm rad^{-1}]	50	(fixed)

Table 3 Results from the parameter estimation, manual adjustment (*) or adjustment according to the rules in (McRuer and Jex, 1967) (**), experiment 2

model	parameter		value
pilot	K_p		0.487
equalisation	Δt_p	[s]	0.213
	τ_{p_L}	[s]	0.5**
	τ_{p_I}	[s]	0**
neural	K_n		1.904
feedback	τ_{n_L}	[s]	0.1*
	τ_{n_I}	[s]	0.253
	Δt_n	[s]	0.02*
IR	K_f		2.853
equalisation	τ_{f_L}	[s]	0*
filter	τ_{f_I}	[s]	0.3*

4. CONCLUSIONS

This paper presented a model for the pilot's neuromuscular system that can be used to describe reactions to disturbances applied to the stick/limb combination. It can also be applied in the investigation of non-conventional sticks, i.e. active sticks, as in (Paassen, 1992), designs for control loading simulations, or to study the pilot's control behaviour with conventional side sticks in situations where a disturbing moment on the stick and the arm should be modelled, for example to account for the accelerations in a moving vehicle.

Two experiments have been described to test the model and estimate parameter values. The model can adequately describe the neuromuscular system's response in these tests. However, it is not possible to estimate all the model's parameters. With the experiment to test the neuromuscular hardware, it did not appear possible to estimate the spring constants of the muscle model's series elasticity and of the skin model. The experiment to test the muscle control and neural feedback models provided similar difficulties. It seems an appropriate step to simplify these models.

It should be noted however, that the human neuromuscular system can adapt itself to the task at hand. The parameter values found in these experiment for the neural feedback and the neural control part of the model can only describe the control with side sticks, and in situations similar to those in the experiments. Other situations may require different parameter values, and a simplified model might not suffice for these situations.

The experiments have, so far, only been carried out for the case of roll control with a side stick. The experimental set-up that was used may be modified to investigate roll control. A new flight simulator facility is being built at the Delft University of Technology, SIMONA (International Centre for Research in SImulation, MOtion and NAvigation). The flight controls in this facility may be used to evaluate the model with yokes. A further focus of research may be the adaptation of the neuromuscular system to different manipulator properties.

5. REFERENCES

Hosman, R.J.A.W., J.B. Benard and H. Fourquet (1990). Active and passive side stick controllers in manual aircraft control. In: *Proceedings of the IEEE Conference on Systems Man and Cybernetics*. Los Angeles.

Kleinman, D.L., S. Baron and W.H. Levison (1970). An optimal control model of human response part I: Theory and validation. *Automatica* **6**, 357–369.

McRuer, D.T. and H.R. Jex (1967). A review of quasi-linear pilot models. *IEEE Transactions on Human Factors in Electronics* **8**, 231–249.

Paassen, M.M. van (1991). Human operator response to visual and proprioceptive inputs in a control task with a side stick. In: *10th European Annual Conference on Human Decision Making and Manual Control*. Liège, Belgium.

Paassen, M.M. van (1992). Biomechanics in aircraft control. In: *Proceedings of the 5th IFAC Symposium on Analysis, Design and Evaluation of Man-Machine Systems*. The Hague.

Paassen, M.M. van (1994). Biophysics in Aircraft Control - A Model of the Neuromuscular system of the pilot's arm. PhD thesis. Delft University of Technology, Faculty of Aerospace Engineering.

Winters, J.M. and L.W. Stark (1985). Analysis of fundamental human movement patterns through the use of in-depth antagonistic muscle models. *IEEE Transactions on Biomedical Engineering*.

APPLYING VIRTUAL REALITY TO DIAGNOSIS AND THERAPY OF SENSORIMOTOR DISTURBANCES

T. KUHLEN*, K.-F. KRAISS*, M. SCHMITT* and C. DOHLE**

**Aachen Technical University, Institute of Technical Computer Science, D-52074 Aachen, Germany*
E-mail: kuhlen@techinfo.rwth-aachen.de

***Heinrich Heine University Düsseldorf, Department of Neurology, D-40225 Düsseldorf, Germany*

Abstract. Lesions of human cortex areas often lead to elementary as well as higher-order disturbances of motor behaviour. In a cooperation between the Institute of Technical Computer Science in Aachen and the Department of Neurology in Düsseldorf, a system based on Virtual Reality (VR) techniques is being developed for the diagnosis of patients suffering from such sensorimotor disturbances. The use of VR comprises two aspects: firstly, VR can be used as an advanced visualization tool for motion analysis by the physician. Secondly, patients can be brought into virtual scenarios where they can perform specific motor tasks for diagnosis and therapy purposes.

Key Words. Medical applications; diagnosis; arm movements; motion parameters; virtual reality; stereo vision; computer graphics; human/machine interaction

1. INTRODUCTION

The term Virtual Reality (VR) describes a computer-generated scenario of objects (virtual environment, virtual world) with which the user (here patient or physician) can interact. In contrast to conventional human-computer interfaces the interaction is designed in three dimensions rather than in two. The combination of three-dimensional computer graphics, special display techniques (head-mounted displays or stereo glasses) and specific input devices (spaceball, dataglove, etc.) allows intuitive manipulation of objects in the virtual world, thus giving the user the impression of being part of the scenario.

This paper introduces a Motor Performance Analysis System (MPAS) being developed in cooperation between the Institute of Technical Computer Science at Aachen Technical University and the Department of Neurology at Heinrich Heine University Düsseldorf, employing Virtual Reality techniques in order to improve the diagnosis of sensorimotor disturbances. MPAS is based on a Silicon Graphics Iris Indigo^2 Extreme UNIX-workstation, which provides the computing and graphics performance necessary for the generation and presentation of VR scenarios.

In principle, sensorimotor disturbances can be divided into elementary dysfunctions like pareses and higher-order disturbances, one of which is apraxia (Freund, 1992). An apraxic patient has lost the capability to spatially and/or chronologically organize and coordinate movement sequences. Apraxias can arise from lesions of the human cortex and are often accompanied by elementary sensorimotor disturbances.

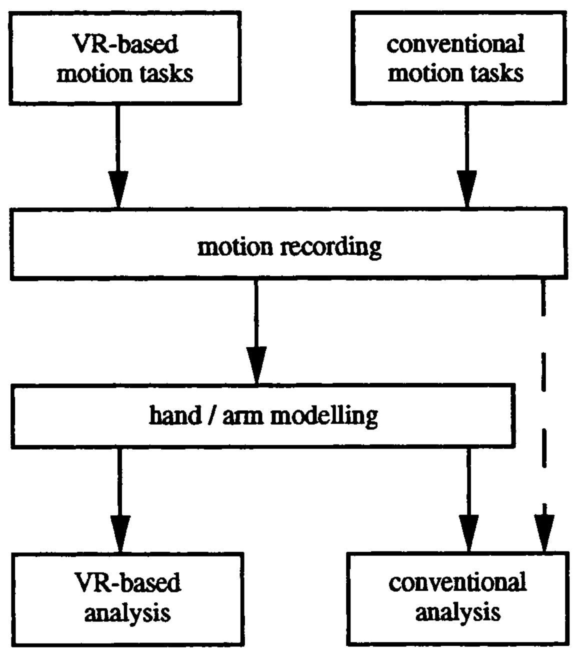


Fig. 1. System overview

In MPAS the use of VR in the field of sensorimotor disturbances comprises two aspects (see Fig. 1): on the one hand, patients can be brought into virtual scenarios where they perform specific tasks

– either to measure their performance for diagnosis (e.g. duration or accuracy of movements) or to train certain actions for rehabilitation purposes. On the other hand, the physician can benefit from three-dimensional visualization capabilities during motion analysis. Both aspects are integrated into MPAS together with existing "conventional" motion tasks and analysis methods.

Sections 2 and 3 briefly discuss the motion recording methods implemented in MPAS and hand/arm modelling respectively. Section 4 refers to the motion analysis component, whereas section 5 deals with the motion task generation component of the system.

2. MOTION RECORDING

To this day, diagnosis of sensorimotor disturbances is essentially based on the diagnostician's experience and is usually restricted to qualitative descriptions of the patient's motor performances (Goldenberg, 1993). Quantitative descriptions of complex motor performances have only recently been used in clinical studies. Since higher-order disturbances resulting from lesions of the cortex affect coordination of the whole upper extremity, measurements cannot be limited to single joints (Mai *et al.*, 1993). Instead, position tracking systems are often used with cameras which record the positions of markers attached to the patient's body (Poizner *et al.*, 1990; Müller *et al.*, 1991).

MPAS uses the Selspot II optoelectronic tracking system. Here, markers are infrared light emitting diodes (LEDs) whose three-dimensional coordinates are calculated from different viewpoints of two or three cameras at a sample rate of up to 100 Hz. The LEDs are illuminated sequentially in order to provide identification in real time. This real time capability makes it possible to utilize measured (tracked) movements not only off-line for later diagnosis by the physician, but also for on-line interaction with objects in a virtual scenario using one's own hand(s).

In addition to Selspot II, MPAS has been equipped with a dataglove from VPL Research in combination with an electromagnetic tracking system from Polhemus Inc. attached to the patient's wrist. However, since those systems are still lacking sufficient accuracy and speed, they are not suited for a precise assessment of motor deficits.

3. HAND/ARM MODELLING

Up to now, modelling techniques are not used for the motion analysis of sensorimotor disturbances (see dotted arrow in Fig. 1). However, an adequate hand/arm model cannot only provide data for existing analysis methods, it is also compulsory in order to realize a three-dimensional animation of motions in a VR-based analysis.

In general, different available recording techniques provide different information, describing the state of the upper extremities only incompletely. An optoelectronic system, e.g., provides three-dimensional trajectories of markers attached to the patient's body, whereas a dataglove yields angle information about the finger joints. Electromagnetic tracking systems yield position and orientation of their sensors.

It is obviously desirable to have a uniform data format that describes as many state variables of hand(s) and arm(s) as exactly as possible for later use in motion analysis. As a consequence, an adequate modelling technique including anthropometric data should be used in order to compute lacking state values from the values supplied directly by the respective motion recording technique.

Using Selspot II, the system has to determine all joint angles from three-dimensional marker positions at any time. For that purpose, a set of LED positions attached to the patient's body has been established, providing sufficient information for angle calculation. Apart from shoulder end elbow markers, for example, at least two LEDs are necessary to provide forearm position and rotation. As an increase of LED number results in a decrease in maximum sampling rate, the LED set is optimized with regard to the number of LEDs.

In MPAS the modelling procedure itself combines inverse kinematics and simple vector calculations (Schmitt, 1994). Since this paper concentrates on Virtual Reality features, the modelling procedure is not discussed in detail here.

4. MOTION ANALYSIS

Currently, motion visualization and analysis in the field of sensorimotor disturbances are not sufficiently supported by contemporary computer techniques. Recorded trajectories supplied by optoelectronic systems are usually simply visualized as two-dimensional plots. Stick diagrams, connecting marker positions at corresponding times by lines, are often used to demonstrate a patient's movement. Values of single joint angles or kinematic parameters such as velocity and acceleration are calculated and presented as two-dimensional diagrams or tables (Winter, 1990).

In contrast, Virtual Reality techniques allow three-dimensional visualization and manipulation of medical data which can be subsequently explored interactively by a physician. In case of motion analysis, the patient's movements are visualized in three dimensions, and the physician changes view position, view direction and view angle in real time.

Basically, trajectories of markers can be visualized in this way. However, more important than this static form of visualization is the possibility to animate a patient's movement using a geometric representation of the hand/arm model as described in section 3. This intuitive way of presentation enables the physician to concentrate on the specific aspects of a motor deficit. Motions can be animated within VR at different rates, thus allowing precise assessment and interpretation. Fig. 2 shows such an animation including the trajectories of markers at thumb and index finger.

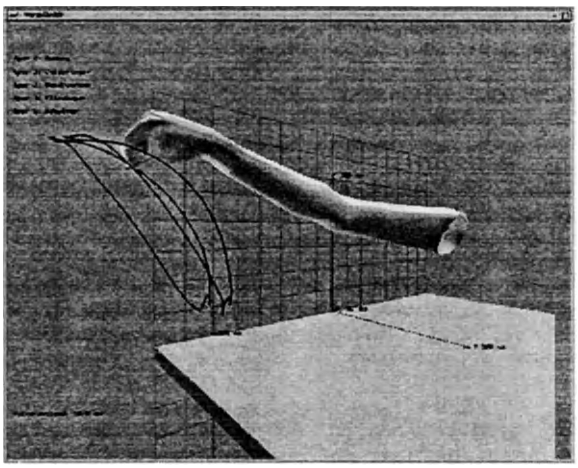


Fig. 2. Screen snapshot from an animation showing a patient's hand movement

In MPAS, three-dimensional visualization is achieved using a high-resolution graphics monitor in combination with stereo shutter glasses (Fig. 3). The monitor alternately displays images to the left and the right eye, while the glasses synchroneously darken the right and left lens respectively. The spaceball, a joystick-like device for six degrees of freedom (translation and rotation, each in three dimensions), enables the physician to move within the virtual scenario and look at the trajectories or the animation from different viewpoints.

5. MOTION TASKS

A usual examination of sensorimotor disturbances employs a pool of functional tests. Patients have to perform motor tasks, which are then analyzed for diagnostic purposes and therapy planning. Motions can be performed on request or as imitations and may include interaction with one or more objects.

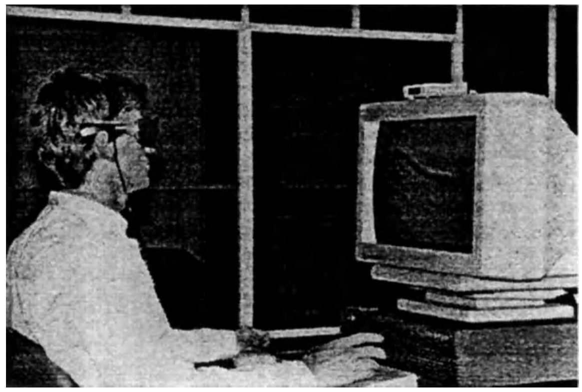


Fig. 3. Physician using the motion analysis component of MPAS

No motion tasks have been developed so far, however, requiring patients to act in Virtual Reality. In many cases the use of virtual instead of real test environments has significant advantages:

Firstly, VR-based tests can be accurately graded, realizing a smooth transition from simple, low-dimensional to more complex motor tasks. Thus, series of motor tasks can be generated not only for diagnostic, but also for therapy purposes. Here, an additional motivating effect for patients arises from the precise feedback in real time.

Secondly, motion tasks of special interest can be shaped that require test environments which are difficult or even impossible to generate in reality. For instance, prehension movements are frequently used as model paradigm for the interaction between transport and manipulation movements in planning and execution. Here, the patient is instructed to reach for and grasp an object, e.g. a cube, as fast and as accurately as possible. New insights about mental information processing can be gained if the object changes position or size during the prehension task. Such effects are difficult to realize using real objects (Paulignan *et al.*, 1991; Gentilucci *et al.*, 1992; Castiello *et al.*, 1993), whereas scaling and moving of virtual, computergraphic objects is trivial.

Futhermore, VR-based motor tasks should lead to a more precise assessment of higher-order disturbances like apraxias. Here, the problem is that apraxias are often accompanied by elementary motor disturbances. In such cases, up to now single testing of the underlying "mental motor images" has not been possible, as the execution requires the use of the patient's own motor system. However, providing different input technologies (3D tracking devices, spaceball, computer mouse) to control complex movements within virtual environments, it will be possible to separate different levels of movement planning and execution.

Visualization concept. A very important question refers to how virtual scenarios should be presented to the patients for the execution of motor tasks. In Virtual Reality systems a head-mounted display is often used, which allows full immersion into a virtual world. However, from the authors' experience, even healthy persons are frequently irritated by the complete loss of visual contact to the real environment. In particular, in full immersive environments the patients' hands have to be represented by computergraphic counterparts. Therefore, a non-immersive visualization technique is preferable, especially with regard to those motion tasks, in which patients have to interact with objects using their own hands.

As a consequence, in MPAS a visualization technique based on a high-resolution graphics monitor and stereo glasses is not only used during motion analysis (see section 4), but also for motion tasks. Since virtual objects can be visualized stereoscopically by this technique, they seem to be situated in free space, i.e. in front of the monitor, and can be grasped as shown in Fig. 4.

However, the interaction with virtual objects also requires that they are perceived by the subject as stationary. In other words, the virtual scenario must be visualized in a way that it is perceived exactly like a scenario of real objects. Therefore the visualization of a virtual scenario has to be adapted dynamically to the patient's view position and direction. The inclusion of these variables into the projection of the current frame in real time, known as "Virtual Holography" or "Fishtank Virtual Reality", is regarded by the authors as essential for the implementation of motor tasks. At present, an electromagnetic tracking system is used in MPAS to measure the patients' head position and orientation, but it is planned to replace it with a video-based tracking system.

At the Institute of Technical Computer Science in Aachen, algorithms have been developed which compensate for non-linear distortions that result from curvature and refraction index of the screen and cause projection errors depending to a great extent on the current viewpoint (Walter, 1994). Figures 5 and 6 may give an impression of the accuracy achievable by these algorithms. While the lower cube is real, the upper one is a holographically visualized virtual cube.

Since Virtual Holography increases the 3D-effect dramatically, the physician can also profit from this technique during motion analyis, even if virtual objects have not to be manipulated here.

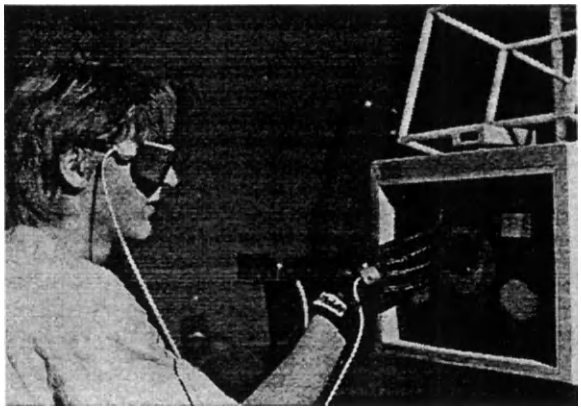


Fig. 4. Interaction with a virtual object, here using a dataglove

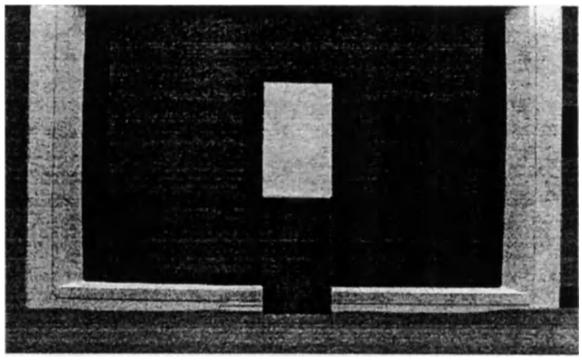


Fig. 5. Real (lower) and virtual (upper) cube as seen by an observer being centered in front of the screen

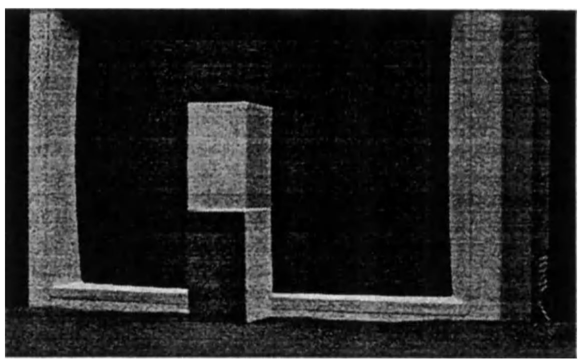


Fig. 6. Moving the head to the right, the projection is changed so that the view to the virtual cube is exactly the same as to the real cube.

6. CONCLUSION

It has been shown that Virtual Reality techniques can be usefully applied in the field of sensorimotor disturbances. Virtual scenarios can be created, in which patients perform specific motor tasks. Using an optoelectronic position tracking system for motion recording as well as for interaction with virtual objects, and using Virtual Holography as an adequate visualization technique, the authors are now developing and testing suitable scenarios. Furthermore, an advanced visualization and animation tool using Virtual Reality technology for diagnosis purposes and therapy planning has been developed, which the physician can profit from during motion analysis.

7. REFERENCES

Castiello, U., K. M. B. Bennett and G. E. Stelmach (1993). Reach to grasp: the natural response to perturbation of object size. *Experimental Brain Research*, **94**, pp. 163 - 178. Springer-Verlag, Berlin.

Freund, H.-J. (1992). The apraxias. In: *Diseases of the Nervous System, Clinical Neurobiology* (A. K. Ashbury, G. M. McKhann and W. I. McDonald, Eds.), 2nd ed., pp. 751 – 767. Saunders, Philadelphia.

Gentilucci, M., S. Chieffi, M. Scarpa and U. Castiello (1992). Temporal coupling between transport and grasp components during prehension movements: effects of visual perturbation. *Behavioural Brain Research*, **47**, pp. 71 - 82. Elsevier Science, Oxford.

Goldenberg, G. (1993). Praxie. In: *Neuropsychologische Diagnostik* (D. Y. von Cramon, N. Mai and W. Ziegler, Eds.), pp. 259 – 286. VCH Verlagsgesellschaft, Weinheim.

Mai, N., M. Blaut and J. Hermsdörfer (1993). Handfunktionen. In: *Neuropsychologische Diagnostik* (D. Y. von Cramon, N. Mai and W. Ziegler, Eds.), pp. 225 - 257. VCH Verlagsgesellschaft, Weinheim.

Müller, F., E. Kunesch, F. Binkofski and H.-J. Freund (1991). Residual sensorimotor functions in a patient after right-sided hemispherectomy. *Neuropsychologia*, Vol. 29, **2**, pp. 125 - 145. Pergamon Press, Oxford.

Paulignan, Y., C. MacKenzie, R. Marteniuk and M. Jeannerod (1992). Selective perturbation of visual input during prehension movements. *Experimental Brain Research*, **83**, pp. 502 - 512. Springer-Verlag, Berlin.

Poizner, H., L. Mack, M. Verfaellie, L. J. G. Rothi and K. M. Heilman (1990). Three-dimensional computergraphic analysis of apraxia. *Brain*, **113**, pp. 85 – 101. Oxford University Press.

Schmitt, M. (1994). Entwicklung eines Bewegungsanalyse-Systems zur Diagnose sensomotorischer Störungen unter Verwendung interaktiver 3D-Computergrafik. Diploma Thesis. Institute of Technical Computer Science, Aachen Technical University.

Walter, P. (1994). Erzeugung von Stereowahrnehmungen durch Virtuelle Holographie mit Anwendung auf die Therapie motorischer Störungen. Diploma Thesis. Institute of Technical Computer Science, Aachen Technical University.

Winter, D. A. (1990). Three-dimensional computergraphic analysis of apraxia. *Biomechanics and motor control of humam movement*, 2nd ed.. John Wiley and Sons, New York.

STUDY OF HUMAN OPERATION OF A POWER DRILL

Denis Rancourt* and Neville Hogan†

**Department of Mechanical Engineering, Laval University, Ste-Foy, (Qué), Canada, G1K 7P4 and †Departments of Mechanical Engineering and Brain and Cognitive Sciences, MIT, Cambridge, MA 02139, USA*

Abstract: This research investigated human operation of a hand-held power drill by analyzing patterns in upper limb posture for various drilling conditions: drilling height, drill length and level of axial force exerted on the drill handle. Subjects used similar postures for drilling locations with similar kinematics. In addition, posture was not statistically different for the range of axial forces and drill lengths studied. Results suggest that human do not try to obtain isotropy or maximum manipulability conditions at the hand while executing the task. Minimization of joint torques appears more plausible.

Keywords: Man/machine interaction, Manipulation tasks, Impedance control, Motor control, Ergonomics, Redundant manipulators, Regularization, Kinematics.

INTRODUCTION

This study investigated human operation of a hand-held power drill. The general goal of the research was to investigate the control strategy that humans use in the operation of hand-held power tools. Better understanding of human performance in such manipulation tasks may be helpful for the automation of industrial processes, to improve productivity or to replace human operators in potentially dangerous work situations. Furthermore, it may form the basis for more precise hand-held tool design guidelines to reduce the prevalence of occupational disorders such as Carpal Tunnel Syndrome (CTS) and Hand-Arm Vibration Syndrome (HAVS). The drilling task was selected because it is a quasi-static task and it requires gross hand motor coordination. The drilling task involves the accurate positioning of a tool and the production of a significant force on an inherently unstable object: those issues are common to the operation of other hand-held power tools such as riveters and power screwdrivers.

Drilling is a complex motor coordination task that requires synergistic control of many muscles and articulations of the human body. Drilling is at maximum a 6 DOF task and the upper limb has at least 10 DOF (6 in rotation and 3 in translation by locating the shoulder in 3D space). Such important redundancy may be used advantageously to optimize performance while still producing the required forces and moments on the drill handle. For instance, it has been suggested that humans choose a limb posture such that they achieve a condition of isotropy (Angeles, 1992) or maximum manipulability at the hand (Yoshikawa, 1990).

Upper limb posture may also be chosen such that a particular stiffness field is apparent at the hand. Studies by Mussa-Ivaldi *et al.* (1985) have shown that the human hand has a significant stiffness magnitude (500 N/m) Stiffness may be changed by simultaneous activation of opposing muscle groups but hand stiffness (and, in general, mechanical impedance) is also a strong function of limb posture. In particular, that study indicated that humans were able to change the orientation of hand stiffness only by limb postural variations. A simple model of the drilling task shows that a minimum hand stiffness is required to stabilize the drill, and this minimum depends on the level of axial force exerted on the drill handle, as well as the drill length and the upper limb posture.

Currently, it is not clear whether drill operators select a particular posture nor whether it is related to any control strategies which are commonly used for the control of redundant manipulators. Because posture is involved in several aspects of the drilling task, patterns in drill operator upper limb posture may provide useful information on the control strategy that is used. This paper presents the results of a study that analyzed the upper limb posture selected by drill operators. Since the operator must compensate for static instability, the research focused on the effect of instability on upper limb posture. Drill operator upper limb posture was measured for different drilling heights, different levels of axial force exerted on the handle and different drill lengths.

1. STABILITY ANALYSIS OF DRILLING TASK

Assume a simple model of a drill that can pivot about the contact point without friction (cf. Fig 1). When an operator exerts an axial force $\mathbf{F_a}$ on a drill handle, there is a destabilizing moment **M** induced:

$$M = F_a x, \quad (1)$$

where x is a small displacement of the drill handle. This moment can be compensated by the action of a spring of stiffness K. A minimum bound on K can be computed from the moment balance about the pivot point:

$$KxL \cos\theta - F_a x \geq 0. \quad (2)$$

For small deviations, the equation reduces to:

$$K \geq \frac{F_a}{L}. \quad (3)$$

This equation indicates that the minimum stiffness required to compensate for the instability is a function of the level of axial force exerted and the drill length. Equation (3), plotted in figure 2, indicates that the required stiffness for a reasonable drill length i.e. 20 to 35 cm, is within the physiological range of human hand stiffness. More complex models show that the kinematics of the manipulator that pushes on the drill is also important.

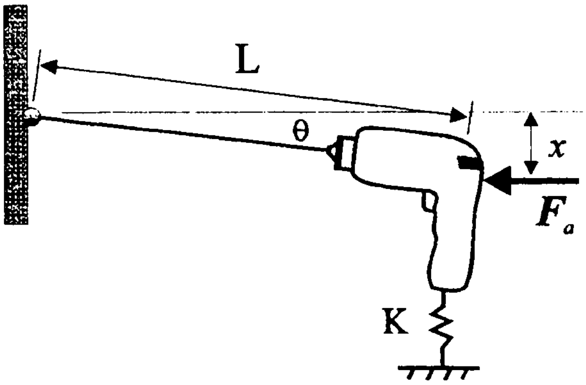


Fig. 1. Mechanical model of instability due to an axial force $\mathbf{F_a}$ exerted on drill handle.

2. METHODS

Subjects were instructed to pick up a drill (1.5 kg, 3/8" Makita power drill model #6510LVR 3.5A), and to drill at given locations on a stainless steel metal plate (cf. Fig. 3). The plate was fixed on a vertical panel mounted on a rolling cart. While the subject drilled, the body posture was recorded both by an infrared SELSPOT camera system and a video camera. Cameras measured the 3D location of infrared LED markers affixed to the upper limb (C7 vertebra, shoulder, arm, elbow, forearm) and the drill.

To avoid lateral instability of the drill bit tip on the metal surface, starter holes were pre-machined into the metal plate at pre-determined locations. The torque transmitted by the metal plate to the drill bit was very small and insensitive to the orientation of the drill because the drill bit was unable to dig into the high strength stainless steel plate. Hence, the drilling conditions did not change even after drilling at the same spot several times.

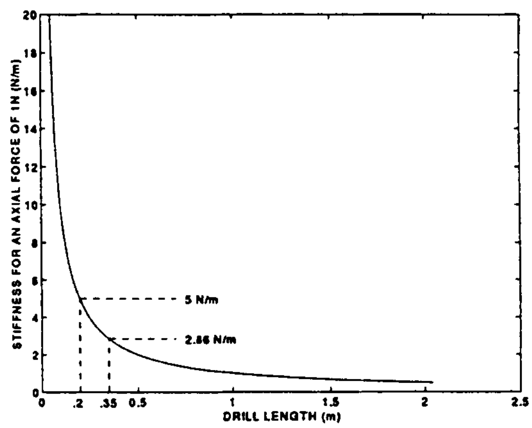


Fig. 2. Normalized transverse linear stiffness variation with drill length. Stiffness values are given for practical range of drill length.

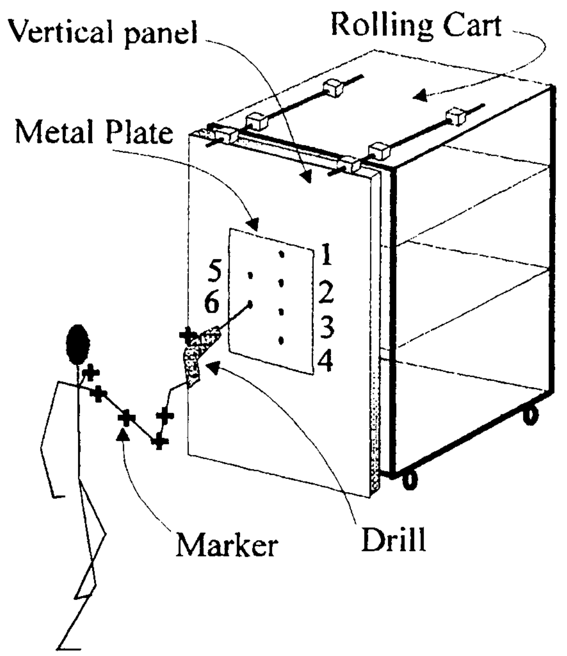


Fig. 3. Experimental setup for drilling experiment.

A Barry-Wright FS6-120A 6 axes force sensor prevented the board from sliding toward the cart while subjects were drilling. Only the axial force perpendicular to the board was measured by the cell (200 Hz sampling rate).

Subjects were allowed to use only one hand to operate the drill and the other hand was not to be used to grasp the cart. Subjects were also not allowed to lean the elbow of their drilling arm against their waist. During drilling, the subject was to maintain a predetermined axial force on the drill handle, typically 65% of maximum exerted force (MEF) and occasionally, 50% or 85% of their MEF. The actual axial force produced was fed back to the subject using a visual display constructed of 5 LEDs. The LED display was fixed at the center of the drilling panel, ensuring that subjects did not modify their body posture to look properly at the LEDs. No additional instructions were given to the subjects.

Each trial consisted of a 2 to 5 second drilling period. Once the subject had stabilized the axial force on the drill (the experiment supervisor had access to a display showing the real time value of the axial force exerted), the infrared camera system acquired kinematic marker data over a .5 second period at 200 Hz sampling frequency. To minimize the influence of previous trials on subsequent ones, between each trial subjects put the drill on a chair and moved away from the panel. Between each trial, the subject was able to rest for 20 seconds to a minute, during which time kinematic data was stored for postprocessing. Resting periods of about 5 minutes were allowed between each set of 3 consecutive tests.

For the first test, called the LOCATION test, subjects were instructed to drill at a randomly selected location from 1 to 6 (cf. Fig. 3). For the second test, called the FORCE test, subjects had to drill at either location 2 or 3, with force levels that randomly varied between 50%, 65% and 85% of their maximum exerting force. For the third test, called the LENGTH test, subjects had to drill at either location 2 or 3, with three different drill bits which resulted in drill lengths of 22, 27 and 32 cm. There were five trials per experimental condition, for a total of 30 trials per test. Two left handed and six right handed subjects participated to the study.

3. RESULTS

3.1 Stick plots

Stick plots were used to depict the subject's upper limb posture. They are line segments that connect the C7 vertebra, the shoulder, the elbow, the hand, and two points on the drill, all computed from measured marker's positions affixed on the subject. A typical example is shown in Figure 4. Those are the results for one subject, all trials of the LOCATION test. If subjects select a posture based only on the kinematics and/or the dynamics of the task, it should be expected that upper limb posture would be repeatable for the same drilling location, and moreover, that posture between locations 2 and 5, or 3 and 6 should be the same. The stick plots support this hypothesis. For instance, the shoulder marker remained within a 6 cm radius from one trial to another. There are some variations of posture which are partly due to the fact that the orientation of the drill was not constrained. Subjects could slightly move the drill up or down by a few degrees without making the drill bit unstable at the surface. The subject-drill system could also rotate in a horizontal plane about the drill tip, without changing the upper limb configuration. These global rotations were removed before statistical analysis.

3.2 Grood & Suntay joint angles

Elimination of global rotations can be done by analyzing posture in the joint angle domain. The rotations occurring at the wrist, the elbow and the shoulder can be described by joint angles which together, define the upper limb posture. To avoid the non-commutativity of Euler angles, the Grood &Suntay (1982) technique was used to determine the Flexion/Extension, External/Internal rotation, and Abduction/Adduction occurring at one particular joint of the upper limb.

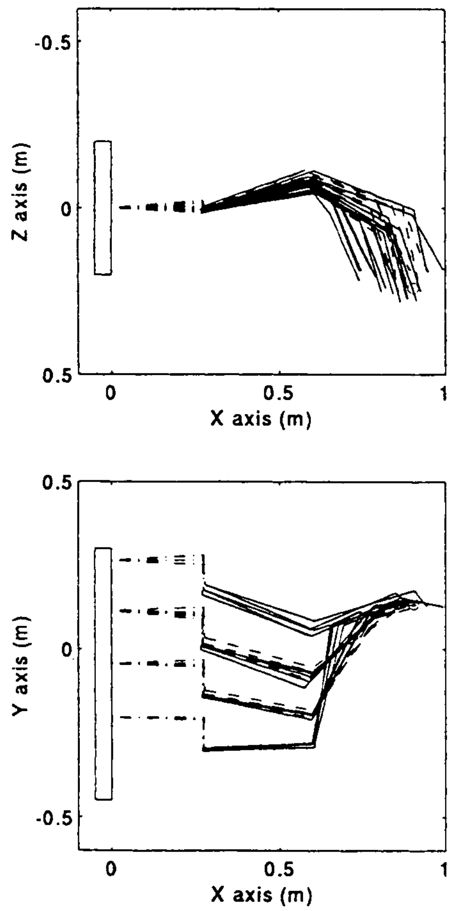


Fig. 4. Stick plots of one subject. Results from 30 trials, LOCATION test. Top plot: top view of the subject; bottom plot: lateral view of the subject. XY plane is vertical, and XZ plane is horizontal. (—) for limb at locations 1 to 4; (--) for limb at locations 5 and 6; (-.-) for drill at any locations. The box represents the vertical board on the rolling cart.

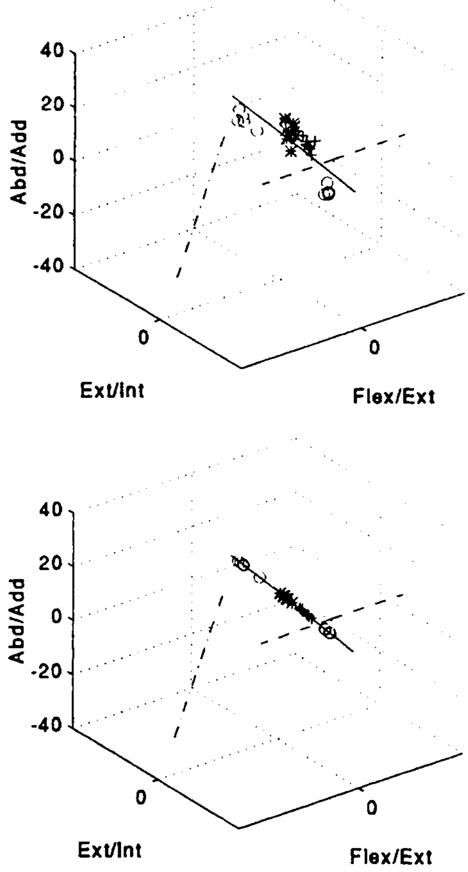


Fig. 5. Grood & Suntay wrist joint angles for one subject. Results from 30 trials, LOCATION test. Top plot (5a): 3D plot of joint angle points. Bottom plot (5b): 3D plot of projected joint angle points on the main eigenvector of the cloud of 30 points. (o) points for locations 1 and 4; (*) points for location 2 and 5; (+) points for location 3 and 6. The straight lines represent the main eigenvectors of the cloud of points for the wrist (—), the elbow (-.-) and the shoulder (--).

The computed wrist joint angles for the LOCATION test of one subject are shown in Figure 5a in a 3D plot. Similar plots can also be made for the elbow and the shoulder. Each symbol represents the state of the wrist joint for one of the 30 trials of the test. The plot shows that points are grouped in separate clusters which are related to any of the six drilling locations. Points for locations 2 and 5, or 3 and 6 are generally overlapped.

Because of the small number of experimental data points, to perform statistical tests, each point was first projected on the main eigenvector of the 30 point cluster. The result is shown in Figure 5b. Then, the distance along the eigenvector between each point and the mean of the points of location 4 were computed and plotted, as illustrated in Figure 6. Based on these distances, called *Linear Angles*, 1D statistical t-tests were performed.

Results for this subject show that the technique used can not tell whether wrist joint angles are different between locations 2 and 5, or 3 and 6 (for either 2%, 5% or 10% significance). However, at any of the three significance levels, results show that posture is different between points 2 and 3, as expected based on the stick plots of Figure 4. Similar results were obtained for the elbow and the shoulder, for five of the eight subjects that were investigated.

The linear joint angle analysis was also used to compare upper limb postures for different force levels and different drill lengths. The analysis showed that the test could not differentiate between postures selected for the different experimental conditions studied.

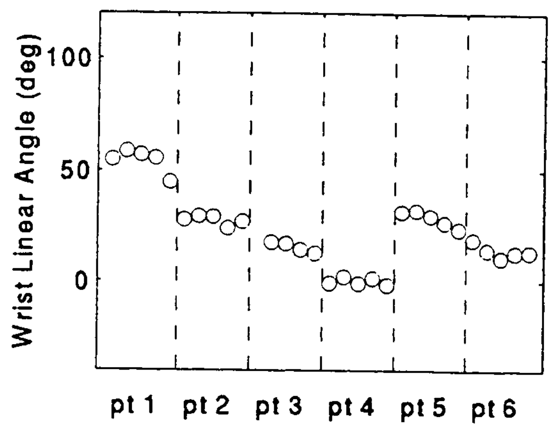


Fig. 6. Wrist linear angles for one subject, LOCATION test.

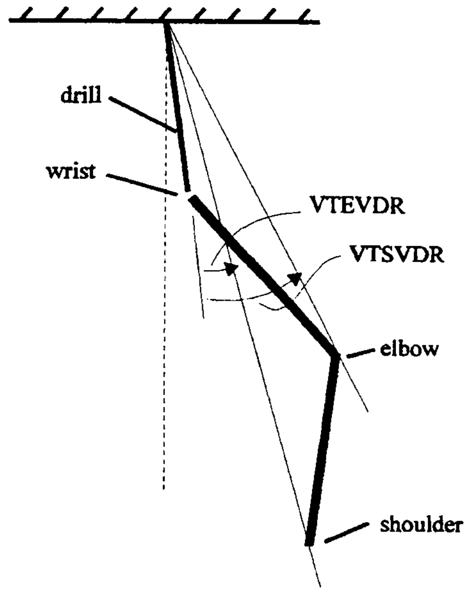


Fig. 7. Illustration of other limb descriptors on a superior view of the subject and the drill.

3.3 *Other upper limb descriptors*

Grood & Suntay linear joint angles have been useful for statistical tests, but they are not physically meaningful. Other limb descriptors are more appropriate such as: (1) the angle between the vector which connects the drill tip to the shoulder, called VTSVDR angle (cf. Figs. 7 and 8), (2) the angle between the vector which connects the drill tip to the elbow, called VTEVDR angle (cf. Figs. 7 and 9), (3) the elbow flexion angle (cf. Table 1), and (4) the distance between the hand and the shoulder (cf. Table 2). Figure 8 shows that most of the right handed subjects (s1 to s6) had VTSVDR angles around 0 degree. This means that the drill axis was aligned with the shoulder, when looking at a superior view. In contrast, left handed subjects (s7 and s8) generally had lower VTSVDR values. Instead, as illustrated in Figure 9, they aligned the elbow with the drill axis since their VTEVDR angle was around 0 degree.

Elbow flexion angles, listed in Table 1, varied across subjects but remained generally constant for any of the experimental conditions studied. The distance between the hand and the shoulder was also generally constant within a subject but varied across subjects.

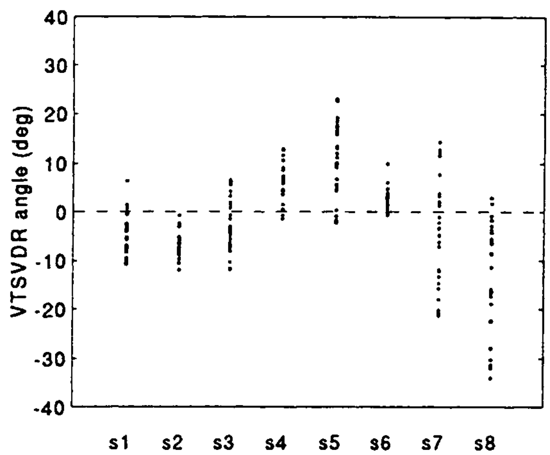


Fig. 8. Angle between the vector which connects the drill tip to the shoulder (VTSVDR angle) for all subjects, LOCATION test.

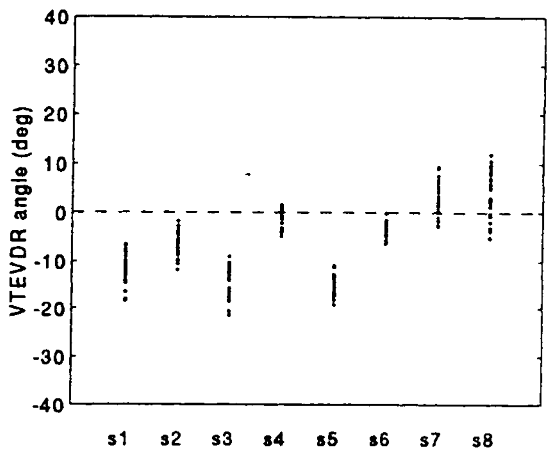


Fig. 9. Angle between the vector which connects the drill tip to the elbow (VTEVDR angle) for all subjects, LOCATION test.

4. CONCLUSIONS

The objective of this research was to investigate whether the analysis of posture would provide useful information on the strategy that humans use to operate a power drill. The small variations of posture for similar drilling conditions suggest that posture is, indeed, controlled and does not vary randomly.

Table 1 Means and standard deviation of elbow flexion angle (EF angle) for all subjects and all tests.

Subject	EF mean	EF std	N
s1	63	10	59
s2	57	16	90
s3	94	7	89
s4	98	7	82
s5	113	11	60
s6	58	8	88
s7	119	6	90
s8	127	6	86

Table 2 Means and standard deviation of hand-shoulder distance (H-S distance) for all subjects and all tests.

Subject	H-S mean	H-S std	N
s1	.55	.03	52
s2	.55	.05	90
s3	.41	.03	89
s4	.43	.03	63
s5	.38	.02	87
s6	.61	.03	88
s7	.31	.02	90
s8	.31	.03	86

The apparent invariance of posture with drill length and axial force exerted suggests that the kinematics of the task is an important factor. However, assuming a 2D two segment model of the upper limb, results show that human operators did not try to achieve isotropy or to maximize manipulability at the hand. These would require elbow flexion of respectively 45 degrees and 90 degrees. Instead, five subjects of eight had elbow flexion over 90 degrees, averaging from 63 to 127 degrees across all subjects.

Variation of limb posture relative to the drill showed that subjects aligned the drill axis with either the elbow or the shoulder joints. This suggests that posture may be chosen to minimize joint torque.

REFERENCES

Angeles, J.(1992) The Design of Isotropic Manipulator Architectures in the Presence of Redundancies. *Int. J. Robotic Research*, Vol. 11, **3**, 196-201.

Grood ES and Suntay WJ (1983). A Joint Coordinate System for the Clinical Description of Three-Dimensional Motions. *J Biomech. Eng.* **105**, 136-144.

Mussa-Ivaldi, F., Hogan, N., Bizzi, E. (1985). Neural, Mechanical, and Geometric Factors Subserving Arm Posture in Humans. J. Neuroscience, Vol. 5, **10**, 2732-2743.

Yoshikawa T (1990) Foundations of Robotics/ Analysis and Control, MIT Press, Cambridge.

ACKNOWLEDGMENTS

This work was performed in the Newman Laboratory at MIT with the support of IRSST, Québec, and NIH grant no. AR-40029.

INTERMITTENCY OF UNIMPAIRED AND AMPUTEE ARM MOVEMENTS

J. DOERINGER* and N. HOGAN**

*Massachusetts Institute of Technology, Department of Mechanical Engineering, Cambridge, Massachusetts, USA

**Massachusetts Institute of Technology, Departments of Mechanical Engineering and Brain and Cognitive Sciences, Cambridge, Massachusetts, USA

Abstract. A significant feature of human arm movements is *intermittency*; the limbs often move as if distinct submovements have been smoothed together. An experiment is outlined which compares a popular arm prosthesis to the unimpaired arm, and a performance measure based on intermittency is introduced. Origins of the intermittency phenomenon are discussed, and potential engineering applications are presented.

Key Words. Arm movements; cybernetics; human factors; human-machine interface; man/machine interaction

1. INTRODUCTION

Although moving the upper limbs is a skill humans execute with almost effortless ease, a good understanding of the schemes by which these limbs are controlled remains elusive. Many decades of serious research have not resulted in robot arms that can duplicate all aspects of human performance; rather, humans have only just begun to understand the characteristics of their motion.

1.1. *Intermittency*

One feature of human movement that seems relatively robust is *intermittency*. The discovery that human limbs often move in an apparently segmented or intermittent fashion was probably first observed by Woodworth (1899). Woodworth noticed that human subjects, when moving a pencil inside a slot to a fixed target, seemed to do so in two steps. The subjects appeared to first move the pencil quickly to a region near the target, and then move the pencil much more slowly in the final approach. Woodworth called these two phases "initial adjustment" and "current control". Much later, models based on intermittency phenomena successfully described other fixed-target movement paradigms (e.g. when target size, movement distance, or movement time was constrained) (Crossman and Goodeve, 1983; Meyer *et al.*, 1988; Meyer *et al.*, 1990).

Intermittency is not limited to pointing to fixed targets; it was also been noted in continuous tracking (in the context of gun positioning) during and after the second world war (Craik, 1947; Craik, 1948). Navas and Stark (1968) determined that the intermittency observed in visuo-motor tracking was not clock-synchronized but rather input-synchronized; after a single stimulus input, subjects could respond to a second at virtually any time after a fixed interval.

Other evidence for intermittency has been observed in circle-tracing (Abend *et al.*, 1982), peg insertion tasks (Milner and Ijaz, 1990), and in cat forelimb isometric force production (Ghez and Vicario, 1978).

If intermittency is fundamental to human movement, the phenomenon should affect the performance of tools controlled by humans. One such tool that has a particularly intimate relationship to the user is an upper-limb prosthesis. Interestingly, arm prostheses possess a relatively low success rate relative to leg prostheses. Although much research has gone into developing new prosthesis models, only about 50% of the amputee population wear the devices. Of those that do, approximately 90% wear the most technologically primitive model, the body-powered type.

1.2. *The Body-Powered Arm Prosthesis*

In this context, "body-powered" means that the power to operate the prosthesis comes from the amputee's own body; the amputee wears a harness that translates shoulder motion (biscapular abduction and upper arm flexion) into elbow flexion (see Figure 1). The harness pulls a cable that lifts the prosthesis forearm; gravity is needed to pull the forearm down and extend the elbow. A second cable serves to lock and unlock the elbow.

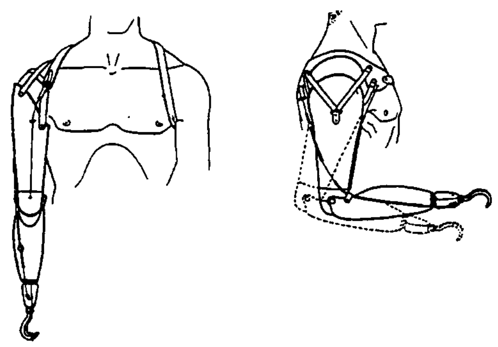


Fig. 1. The body-powered prosthesis.

The alternative to body-powered prostheses is one of several myoelectric models. Myoelectric arm prostheses measure the small electric signals from the amputee's remaining musculature (usually the stump muscles) as control inputs. The prosthesis moves via one or more electric motors and battery pack. Of course, exactly how the device responds to the myoelectric signals depends on how it was designed and/or programmed.

Myoelectric arm prostheses were developed in order to give amputees a more natural man/machine interface—an approximation to an ideal *cybernetic linkage*, in which the input and output connections of the artificial arm are identical to those of the real arm.

Many researchers have proposed that the relative low popularity of myoelectric arms is because they do not allow amputees the haptic sensation of movement; there is no direct output connection from the prosthesis to the neural system. Some propose that intact joints should be used as control sites (rather than isometric muscle activity); the state of each prosthesis degree of freedom should be a unique function of an intact joint degree of freedom. Prostheses controlled this way are referred to as "Extended Physiological Proprioception" (EPP) devices; it has been speculated that a reason for the success of body-powered prostheses is that they facilitate EPP.

Of course, the body-powered prosthesis does have distinct disadvantages. The relatively small ratio of shoulder displacement to prosthesis forearm displacement results in a small load capacity and high harness forces. These forces are typically exerted on the axilla, which is ill-suited for load bearing.

The overwhelming popularity of the body-powered prostheses despite its disadvantages led

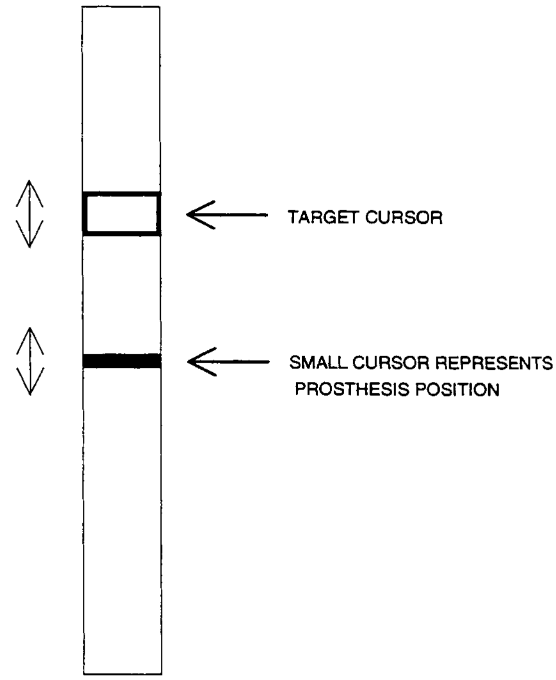


Fig. 2. Display used for all tasks.

us to perform several experiments with experienced body-powered users. The purpose of these experiments was to explore different performance measures in comparing body-powered prosthesis movements to movements of the unimpaired arm. This paper will briefly present some results of one experiment.

2. A "CLASSIC" EXPERIMENT TO MEASURE ARM PERFORMANCE

Five male users of body-powered prostheses performed a series of simple pursuit tracking tasks on both the prosthesis and unimpaired sides. All subjects had used a body-powered prosthesis for at least six months, and had no physical or mental disabilities other than the loss of one arm.

2.1. *Methods*

A potentiometer-based electrogoniometer was fastened to either the unimpaired or prosthesis arm of the subject, depending on which side was being tested. Subjects stood with their arms at their sides; they were told to look at a computer workstation display (see Figure 2). The display consisted of a hollow box (the target) and a solid bar (representing the position of the forearm), both of which moved up and down inside a narrow rectangle. The subjects were instructed to keep the bar inside the box.

Target motion corresponded to bandwidth-limited quasi-random Gaussian signals (standard deviation 14.6 degrees) whose amplitude covered approximately half the range of prosthesis elbow motion (centered at 67.5 degrees—the approximate middle of elbow travel). 10th order unity gain low pass Butterworth filters were used on Gaussian noise to generate four bandwidths: 0.16, 0.40, 0.96, and 1.84 Hertz. The signals ranged in length from 200 to 400 seconds. Subjects tracked two signals at each bandwidth.

2.2. *Results and Discussion*

Figure 3 shows some typical data after smoothing. For clarity, only 15 seconds of the trial is shown. Note that the response to the smoothly moving target is a series of short, quick movements, typical of the intermittency phenomena observed by other researchers. The short movement segments are much more visible in the velocity traces than in the position traces. On the unimpaired side, the movements seem to occur with greater frequency and appear to overlap. On the prosthesis side, the movements are much more separate and distinct.

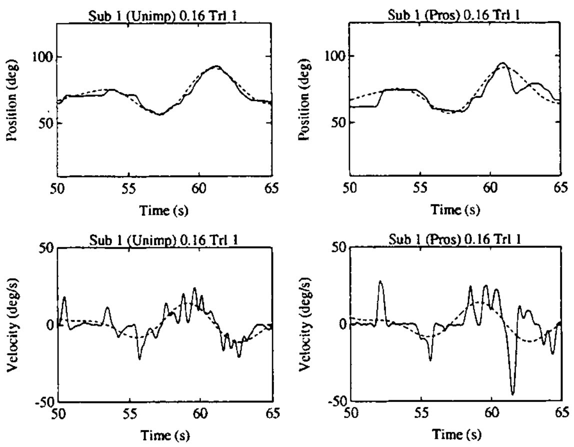


Fig. 3. Example of tracking data (Subject 1, 0.16 Hertz cutoff). Dotted lines represent the target position and velocity, while solid lines represent position and velocity of the subject's unimpaired elbow (left plots) and body-powered prosthesis (right plots).

A simple way to compare performance between the unimpaired and prosthesis sides is to calculate mean squared error between target and cursor. Results for all subjects and all trials can be seen in Figure 4. Note that although mean squared error increases with bandwidth (as expected), it is decidedly unclear as to which side is superior at high bandwidths. The relative performance of Subjects 1, 2, and 5 grows and then shrinks as bandwidth increases, while the performance of Subjects 3 and 4 seems to continually increase.

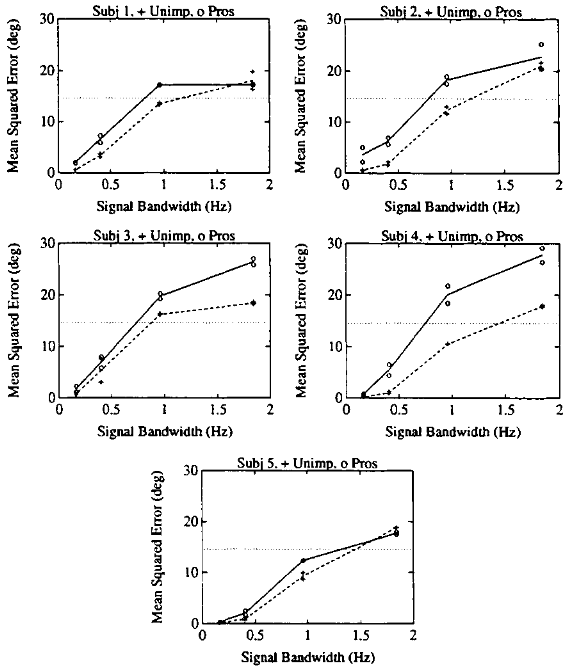


Fig. 4. Mean squared error of all tracking trials. Mean squared error is displayed as a function of tracking signal bandwidth. Plus signs and dotted lines represent the unimpaired side, while circles and solid lines represent the prosthesis side. The horizontal line represents the standard deviation of the input signals.

In contrast to the mean squared error measure, a performance measure based on the intermittency phenomenon produces clearer results. Figure 5 shows the results of counting and measuring the peaks in the velocity data for each trial. Average peak rates for the first 90 seconds of each trial are plotted versus target signal bandwidth.

Note that the subjects do not produce discrete actions at constant rates across trials. For most subjects, the prosthesis can keep up with the unimpaired arm at low frequencies, but as signal bandwidth increases, the prosthesis cannot keep pace with the unimpaired arm. The data also suggest that velocity peak production rate may asymptotically be approaching a limit (somewhere over 3 Hertz), but further experimentation would be necessary to determine this limit conclusively.

3. CAUSES OF INTERMITTENT BEHAVIOR

The causes of intermittent movement behavior are not well agreed upon. Researchers have noted that in the visual tracking paradigm, intermittency can be reduced (but not removed completely) by temporarily removing visual feedback while tracking a repeating waveform (Miall *et al.*, 1993). In the compensatory tracking paradigm (where only error is displayed) intermittency can be altered by

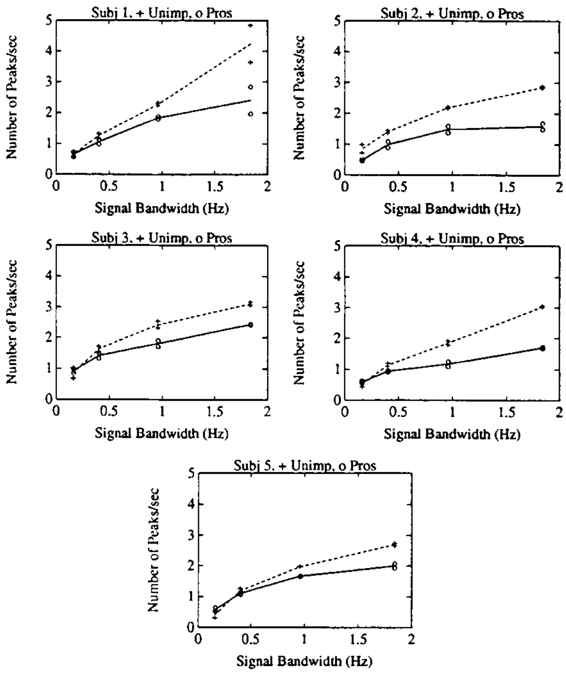


Fig. 5. Summary of tracking velocity peak data. Peak rates are plotted versus target signal bandwidth. Plus signs and dotted lines represent the unimpaired side, while circles and solid lines represent the prosthesis side.

changing the apparent error "deadzone" (the size of the acceptable target region) (Wolpert *et al.*, 1992). It has also been pointed out (Milner and Ijaz, 1990) that the times between intermittent actions in reaching (150-200 ms) are similar to estimates of visual reaction time obtained from humans responding to target shifts (Georgopoulos *et al.*, 1983). In our lab, we are currently pursuing a series of non-visual tracking experiments to be performed alongside analogous visual tracking tasks to quantify the visual dependence of the intermittency phenomenon.

Yet another speculation about intermittency is that it may be related to task learning (Milner and Ijaz, 1990). Intermittent actions may be part of an overall learning strategy in which the actions are the basic constructs of a complex motor command. A first approximation to a desired movement could be made by using a series of intermittent actions, with relatively high-level brain functions checking the success at the end of each action. With practice, these submovements could be blended together to form larger, more complex submovements. Once the desired task was completely learned, the high-level part of the brain would only need to "download" the completely learned movement at the beginning of the task; the relatively expensive high-level control would then be more available for other functions. Experiments are currently ongoing in our laboratory to test this hypothesis.

Intermittency may prove useful as an approach to partially constrained motion problems. Russell (1990) found preliminary evidence that a segmentation strategy may be used in a crank turning task. Our current work involves further experimentation related to these results.

4. UTILITY OF INTERMITTENCY

The simple experiment described in this paper showed how a performance measure based on intermittency could potentially reveal more information than one based directly on the task objective. Although the purpose of the task was to "keep the bar inside the box", the subject strategy was to execute intermittent corrections at high output rates, and these rates were noticeably distorted on the prosthesis side. If the intermittency phenomenon is sufficiently robust, features like velocity peak height, shape, number, and rate could feasibly be used to quantify the performance of many free motion tasks with many different kinds of tools. Since intermittent submovements appear to be a "normal" feature of human motor behavior, the phenomenon could provide a quantitative measure of motor normalcy in, say, humans with neuromotor deficiency. Such a measure would then be extremely useful in identifying therapies which work and designing machines for effective rehabilitation.

The intermittency phenomenon may be a result of feedback to a set of memorized "waveforms" (Miall *et al.*, 1993). If this is indeed the case, then a relatively small set of waveforms could feasibly be used to describe extremely complex human movements. If such a movement "alphabet" could be identified, the transmission of human movement information in telemanipulation and/or remote virtual environment applications could be made much easier. Related to reduced information transmission is the potential to use the movement alphabet to make bounded error predictions. These predictions would be useful to offset measurement and transmission delays in the aforementioned applications.

If intermittent submovements are truly used in conjunction with human learning, the idea should be transferable to autonomous robots learning novel tasks. Consider a robot with a limb controlled in a hierarchical fashion; a high-level controller calculates appropriate movement trajectories that are sent down to a simple low-level controller (which keeps the limb stable). The high-level controller could pick from a library of potential submovements in order to complete a movement task slowly and carefully; it would only need to "pay attention" when a selecting a new submovement. After the entire task was accomplished once, the robot could try again, this time blending the most successful submovements together. When the robot had fully learned the task, the high-level controller would only provide computation at the beginning of the complete movement, thus freeing computational resources during the movement for other machine functions.

5. CONCLUSION

Intermittency is a feature of human movement that has been observed in many separate situations, but much work still needs to be done in exploiting the phenomenon for engineering purposes. A simple experiment comparing unimpaired arms to body-powered arm prostheses was presented, and a performance measure based on the intermittency phenomenon was shown to be useful. The phenomenon could be partially due the delays in visual feedback; it could also be part of a learning strategy or a control method for partially constrained movements. A better understanding of intermittency will help in evaluating human-machine interfaces, transmitting human movement information, and designing autonomous robot controllers.

6. REFERENCES

Abend, W., E. Bizzi and P. Morasso (1982). Human arm trajectory formation. *Brain* **105**, 331–348.

Craik, K.J.W. (1947). Theory of the human operator in control systems. Part 1: The operator as an engineering system. *British Journal of Psychology* **38**, 56–61.

Craik, K.J.W. (1948). Theory of the human operator in control systems. Part 2: Man as an element in control systems. *British Journal of Psychology* **38**, 142–148.

Crossman, E.R.F.W. and P.J. Goodeve (1983). Feedback control of hand movement and Fitts' law. *Quarterly Journal of Experimental Psychology* **35A**, 251–278. Original work presented at the meeting of the Experimental Psychology Society, Oxford, England, July 1963.

Doeringer, J. and N. Hogan (1995). Performance of above elbow body-powered prostheses in visually-guided unconstrained motion tasks. *IEEE Transactions on Biomedical Engineering.* (In Press).

Georgopoulos, A.P., J.F. Kalaska and J.T. Massey (1983). Interruption of motor cortical discharge subserving aimed arm movements. *Experimental Brain Research* **49**, 327–340.

Ghez, C. and D. Vicario (1978). The control of rapid limb movements in the cat. II. Response latency. *Experimental Brain Research* **33**, 173–189.

Meyer, D.E., J.E.K. Smith, S. Kornblum, R.A. Abrams and C.E. Wright (1990). Speed-accuracy tradeoffs in aimed movements: toward a theory of rapid voluntary action. In: *Attention and Performance XIII* (M. Jeannerod, Ed.). Chap. 6, pp. 173–226. Lawrence Erlbaum Associates. Hillsdale, NJ.

Meyer, D.E., R.A. Abrams, S. Kornblum, C.E. Wright and J.E.K. Smith (1988). Optimality in human motor performance: ideal control of rapid aimed movements. *Psychological Review* **95**, 340–370.

Miall, R.C., D.J. Weir and J.F. Stein (1993). Intermittency in human manual tracking tasks. *Journal of Motor Behavior* **25**, 53–63.

Milner, T.E. and M.M. Ijaz (1990). The effect of accuracy constraints on three-dimensional movement kinematics. *Neuroscience* **35**(2), 365–374.

Navas, F. and L. Stark (1968). Sampling or intermittancy in hand control system dynamics. *Biophysical Journal* **8**, 252–302.

Russell, D.L. (1990). An Analysis of Constrained Motions in Manipulation. PhD thesis. Department of Mechanical Engineering, M.I.T.

Wolpert, D.M., R.C. Miall, J.L. Winter and J.F. Stein (1992). Evidence for error deadzone in compensatory tracking. *Journal of Motor Behavior* **24**, 299–308.

Woodworth, R.S. (1899). The accuracy of voluntary movement. *Psychological Review Monograph Supplement* **3**(13), 1–114.

ECOLOGICAL INTERFACE DESIGN: A RESEARCH OVERVIEW

Kim J. Vicente

Cognitive Engineering Laboratory, Department of Industrial Engineering
University of Toronto, 4 Taddle Creek Rd, Toronto, Ontario CANADA M5S 1A4
Email: benfica@ie.utoronto.ca

Designers of advanced computer-based interfaces for complex work environments are facing important design challenges, but research has not kept pace with their needs. Ecological interface design (EID) is a theoretical framework for designing interfaces for complex human-machine systems. This paper provides a review of some of the research conducted on EID. The impact that EID has had on industry in terms of technology transfer is also addressed, as are a set of promising issues for future research.

Keywords: human-machine interface, cognitive systems, fault diagnosis, graphic displays, human factors, nuclear power stations, process control

1. INTRODUCTION

Many industries are transitioning from traditional (e.g., analogue, hard-wired) instrumentation to computer-based technology in designing advanced interfaces for complex human-machine systems. If this trend is to result in an improvement in performance and safety, or at the very least, to maintain current levels where they are, then a deep understanding of the impact that such technology can have on human performance is required. More specifically, one would like to base the design of such advanced interfaces on a set of defensible principles. Such principles must satisfy a number of stringent requirements. First, they must be well suited to the demanding characteristics of complex work environments. Second, they should be based on sound research so that designers and regulatory bodies can be confident that the new designs created from such principles will in fact lead to acceptable levels of performance and safety. Third, such principles should also be perceived to be of practical value to industry, otherwise they will be just an academic curiosity. There is a consensus in the cognitive engineering community that a framework satisfying all of these requirements does not yet exist. This means that designers currently do not have much guidance from the research community as they design advanced interfaces based on information technology.

Ecological interface design (EID) is a theoretical framework for interface design for complex human-machine systems that is intended to satisfy the requirements just outlined. The goal of EID is to allow operators to take advantage of their powerful perception and action capabilities while simultaneously providing the support required for problem solving activities that are required to effectively deal with unanticipated events (Vicente and Rasmussen, 1992). This paper integrates some of the recent research on EID. In the next section, the EID framework will be described in more detail.

2. ECOLOGICAL INTERFACE DESIGN

EID is based on two seminal concepts from cognitive engineering research, the abstraction hierarchy (AH) and the skills, rules, knowledge framework (Rasmussen, 1986). The AH is a multilevel knowledge representation framework that can be used to develop physical and functional plant models, as well as the mappings between them. It is used in

EID to identify the information content and structure of the interface. In addition to specifying interface content, the principles of EID also suggest how information should be displayed in an interface. The idea is to take advantage of operators' powerful pattern recognition and psychomotor abilities. Thus, EID recommends that information be presented in such a way as to promote skill- and rule-based behavior, allowing operators to deal with task demands in a relatively efficient and reliable manner. Knowledge-based behavior is also supported by embedding an AH representation of the work domain in the interface. This provides operators with an external visualization of plant structure and dynamics which offers support during abnormal situations requiring problem solving.

The three principles of EID, each corresponding to a given level of cognitive control, are as follows:

1. Skill-based behaviour (SBB) - to support interaction via time-space signals, the operator should be able to act directly on the display, and the structure of the displayed information should be isomorphic to the part-whole structure of movements.
2. Rule-based behaviour (RBB) - provide a consistent one-to-one mapping between the work domain constraints and the cues or signs provided by the interface.
3. Knowledge-based behaviour (KBB) - represent the work domain in the form of an AH to serve as an externalized mental model that will support knowledge-based problem solving.

For a detailed description and justification of these principles, see Vicente and Rasmussen (1990, 1992). For a detailed example showing how these principles can be applied to interface design, see Vicente and Rasmussen (1990).

Because the AH plays such an important role in EID, it is important to describe it in more detail. The AH is a multilevel representation format that describes the various layers of constraint in the work domain. Each level represents a different model of the system. For many complex human-machine systems, five levels of constraint have been found to be of use:

1. The purposes for which the system was designed (Functional Purpose);
2. The intended causal structure of the process in terms of mass, energy, information, or value flows (Abstract Function);
3. The basic functions that the plant is designed to achieve (Generalized Function);
4. The characteristics of the components and the connections between them (Physical Function);
5. The appearance and spatial location of those components (Physical Form).

Higher levels represent *functional* information about system purposes, whereas lower levels represent *physical* information about how those purposes are realized by plant components.

There are two advantages to adopting the AH as a basis for interface design (Vicente and Rasmussen, 1992; Bisantz and Vicente, 1994). First, this approach allows one to identify, a priori, the information needed to cope with events which are unfamiliar to operators and which have not been anticipated by designers. This is a very important property since such unanticipated events pose the greatest threat to system safety in complex industrial systems. Whereas traditional approaches to interface design rarely make an attempt to deal with this problem, the AH was explicitly designed to deal with unanticipated events. Secondly, the AH is also a psychologically relevant problem representation. There is a significant body of empirical research from a number of diverse domains showing that problem solving protocols can be mapped onto an AH representation (see Rasmussen, 1986 and Vicente and Rasmussen, 1992 for reviews). Thus, in addition to satisfying the engineering requirement of containing the information needed to cope with unanticipated events, the AH also satisfies the psychological requirement of providing a representation that is consistent with operators' problem solving processes.

As a result of its unique characteristics, EID is an obvious candidate for designing graphic, computer interfaces for complex human-machine systems where unanticipated events can occur. In the following section, some of the previous research conducted on EID will be reviewed.

3. LITERATURE REVIEW

3.1 Starting Point

Initial work on EID was conducted in the process control domain. A review of the research and development efforts on multilevel interface design for power plant control rooms revealed that several researchers had developed interface designs that were based on an AH representation of the plant. Thus, these previous efforts were of direct relevance to EID. Unfortunately, up until the time that the review was conducted (1990), no experiment had been conducted comparing a multilevel interface based on an AH representation to any other type of interface (Vicente, 1992). As a result, the direct empirical evidence for the EID framework was very weak, not because the results obtained were equivocal but rather because appropriate studies simply had not been conducted. This provided a starting point for a research program on EID.

The remainder of this section will describe some of the recent research that has been conducted on EID.

To reduce the scope to a manageable level, the review will be focused on the research that has been conducted by the Cognitive Engineering Laboratory (CEL) at the University of Toronto. Most of this research has been conducted using a simplified, but representative, thermal-hydraulic process simulation. Thus, the next subsection will begin by describing this system.

3.2 Research Vehicle

Most of the CEL research on EID has been conducted with two related systems, DURESS and DURESS II. DURESS II is an updated, interactive version of DURESS, a thermal-hydraulic process control microworld simulation program. The original DURESS system was a non-interactive system that presented subjects with real-time, "canned" scenarios. Subjects were able to view the behaviour of the system but not to control it. DURESS II, in addition to some minor structural changes, differs from the original primarily in that it allows for real-time interaction so that subjects may actively control the system.

DURESS II consists of two redundant feedwater streams (FWSs) that can be configured to supply water to either, both, or neither of the two reservoirs. Each reservoir has associated with it an externally determined demand for water that can change over time. The system's purposes are twofold: to keep each of the reservoirs at a prescribed temperature (40° C and 20° C), and to satisfy the current mass (water) output demand rates. To accomplish these goals, the subject has control over eight valves, two pumps, and two heaters. All of these components are governed by first order lag dynamics, with a time constant of 15 seconds for the heaters and 5 seconds for the remaining components. A diagram of the system can be found in Vicente and Rasmussen (1990).

Two different interfaces for DURESS II have been designed to date. The first one is based on a piping and instrumentation diagram (P&ID) representation of the system. This interface provides a physical view of the system, focusing on displaying the current state of all of the components and the goal variables. The P&ID format was chosen because it is typical of how existing computer interfaces for process control systems have been designed. Thus, it serves as a meaningful control condition.

The second interface for DURESS II was based on the principles of EID. A detailed explanation of how this interface was designed can be found in Vicente and Rasmussen (1990), so only a brief overview will be provided here. The EID interface contains all the information contained in the P&ID interface, but it also contains higher-order functional information that was identified through an AH analysis of DURESS (see Vicente and Rasmussen, 1990). Examples of functional information include valve flowrates, heat transfer rates, and in particular, mass and energy topologies. This added information provides a basis for supporting effective reasoning during unanticipated events, one of the important goals of the EID framework. In addition to the inclusion of this high-order relational information, another distinctive feature of the EID interface is the form in which information is provided. The interface was designed so as to preserve a one-to-one mapping between domain constraints and the salient, perceptual properties of the interface, thereby providing an external visualization of the system that is intended to enhance information extraction. This is in keeping with the EID goal of supporting the power of perception and action. A more detailed description of the rationale behind the design of this interface can be found in Vicente and Rasmussen (1990).

3.3 Laboratory Findings

The first experimental evaluation of EID was conducted in the context of the non-interactive version of DURESS (Vicente, 1992a). The goal of that study was to compare the earlier versions of the P&ID and EID interfaces in terms of how well they support problem solving behaviour, or KBB. The experimental evidence indicated that the EID interface provided better support for KBB than the P&ID interface. While this finding was encouraging, the study was limited in several ways. First, the subjects were either theoretical experts or novices at generic thermal-hydraulic principles, but neither group had any substantial experience with DURESS itself. Second, the experiment evaluated the subjects' ability to diagnose and remember the values of dynamic, "canned" scenarios. Therefore, subjects did not interactively control the system. While there were reasons for conducting the experiment under this restricted set of conditions (cf. Vicente, 1992a), it remained to be seen whether the advantage of the EID interface would still hold with subjects with extensive experience controlling DURESS.

The second evaluation of EID was also conducted in the context of DURESS (Vicente, Christoffersen, and Pereklita, in press). In that study, subjects again diagnosed real-time canned scenarios, but this time, only using the EID interface. In addition, verbal protocols were collected as the subjects tried to diagnose the nature of the events presented to them. A process tracing analysis was conducted by mapping subjects' verbalizations onto a two dimensional problem space representation of DURESS, defined by an AH and a part-whole hierarchy (Rasmussen, 1986). Process measures of performance were correlated with product measures of performance to determine whether there were any statistically significant relationships between subjects' cognitive strategies and their diagnosis accuracy. The results indicate that the greater the extent to which subjects adopted the top-down "zooming-in" strategy, that the AH is intended to support (Rasmussen, 1986), the more accurate their diagnosis performance was. This experiment was the first to empirically demonstrate the problem solving advantages associated with reasoning in an AH problem space. Nevertheless, the results are limited by the fact that the subjects did not interactively control the system.

A third study was designed to overcome the limitations of the previous two (Pawlak and Vicente, 1994). The experiment was conducted with the interactive DURESS II simulation so that subjects could control the system components. Slightly revised versions of the P&ID and EID interfaces were compared. Subjects were given extensive practice at controlling the system (1 hour per weekday for 4 weeks) with one of the two interfaces before their performance on normal events and unfamiliar faults was evaluated. The results revealed that, under normal conditions, there was no performance difference between the EID and P&ID interfaces. However, dual task performance results indicate that the P&ID interface relies more on verbal resources, whereas the EID interface requires more spatial resources. Furthermore, a process tracing analysis of the fault trials showed that the EID interface led to faster fault detection and more accurate fault diagnosis performance. In addition, two deficiencies of the EID interface were identified, one suggesting a need for integrating trend information with emergent feature displays, and the other suggesting that displays tailored for enhanced perception of the system state may not always be well suited for fault compensation. The primary contribution of this study, then, was to compare the P&ID and EID interfaces under a more representative (Brunswik, 1956) set of experimental conditions.

The most recent experimental evaluation of EID was intended to investigate the long-term influences of EID on operator performance and knowledge (Christoffersen, Hunter, and Vicente, 1994). The EID and P&ID interfaces for DURESS II were again compared. More specifically, a longitudinal experiment lasting 6 months was conducted to compare these interfaces under a variety of conditions, including normal trials, routine faults, and non-routine faults. Just as in the previous study, subjects controlled the system every weekday (not including holidays) for about one hour per day. Product measures (e.g., time, actions) and process measures (e.g., verbal protocols) of performance were collected. In addition, several knowledge elicitation measures were occasionally administered to determine how the subjects' knowledge organization evolved over time. At the end of the experiment, subjects switched interfaces and had to control DURESS II under normal and routine fault conditions using the new interface. The primary findings of the study were threefold. First, on normal trials, there was very little difference in the average performance of the interface groups. The group using the P&ID interface, however, consistently showed more variability in their performance, occasionally taking much longer than usual to complete the required tasks. This effect was found to hold even after 5 months of practice, and was specific to the interface, not to the subjects. Second, for both routine and non-routine faults, the EID interface was found to lead to better performance, especially with respect to diagnosis accuracy. These effects seemed to stem from strategy differences between interface groups, which were in turn a result of the interaction between the subjects' knowledge and the information provided in the interfaces. Third, subjects using the EID interface who actively explored the system and reflected on the feedback provided were able to achieve levels of adaptation and performance not observed with subjects using the P&ID interface. When adopting a surface approach to learning, however, subjects using the EID interface were likely to exhibit a very shallow knowledge base and poor performance (although no worse than that attained by subjects using the P&ID interface with a comparable level of motivation). Thus, it appears that there are certain preconditions which have to be satisfied if the benefits of the EID interface are to be fully enjoyed.

3.4 Industrial Prototype

Although the results obtained from the laboratory studies with DURESS and DURESS II have been very encouraging, it is important to determine if the EID framework can be scaled up to complex industrial systems. As a first step in this direction, a project was undertaken with ABB Corporate Research - Heidelberg to design a prototype EID interface for the feedwater subsystem of an ABB conventional power plant (Dinadis and Vicente, 1994). This proved to be a very valuable experience.

The study showed that the EID framework can be meaningfully applied to industrial systems that are much larger in scale than the DURESS system that had been used to evaluate EID. An AH analysis of the feedwater subsystem was conducted and served to identify the information content of the interface. This information was then mapped onto salient perceptual cues in the interface that are intended to make information extraction accurate and efficient.

Another significant outcome of the feasibility study was a proof of principle that EID can be integrated with other interface design concepts. EID was only intended to address some basic issues in interface design (Vicente and Rasmussen, 1992), and so there are several important design problems for which it does not provide guidance. Perhaps the most salient example is the issue of navigation. There are many different ways to display a multilevel interface based on an AH across display pages or windows (Vicente, 1992c), but EID does not prescribe which is the most effective. In fact, there is very little empirical evidence on which to base such a decision (Vicente, 1992b). However, the concept of visual momentum (Woods, 1984), which was specifically developed to address this issue, was used in the development of the ABB prototype.

Similarly, EID is also silent on specific issues about coding of visual form. Because it is intended to be a general framework (Vicente and Rasmussen, 1992), EID motivates designers to pursue certain objectives (e.g. mapping domain constraints onto perceptual signs in the interface), but it does not offer a set of procedural steps for how to achieve those objectives. Again, however, it was possible to rely on other design principles (i.e., perceptual organization techniques) to address this interface design issue. These examples reinforce the point made by Vicente and Rasmussen (1992) that following the principles of EID alone does not allow one to design an effective interface for large scale systems. However, these results also show that EID provides a solid basis for design that addresses fundamental issues, and moreover, that it can be integrated with other important design principles. This makes EID a viable candidate for designing interfaces for complex industrial systems.

4. TECHNOLOGY TRANSFER TO INDUSTRY

As mentioned in the introduction, if an interface design framework is to be more than an academic curiosity, it must be perceived by industry to be relevant to significant design problems. How well has EID fared in this regard?

Despite the relative novelty of the framework, there has already been a limited success in transferring EID to industry. This technology transfer has occurred primarily in the nuclear and process control industries and to varying degrees. A very modest form of technology transfer has occurred with AECL Research and Honeywell. Both of these companies have adapted the EID interface for DURESS into prototypes that are intended to illustrate state-of-the-art interface design concepts that may find themselves in advanced control rooms of the future.

Researchers at the Nuclear Engineering Laboratory of Toshiba in Japan (Monta *et al.*, 1991) have gone much further. They have adopted EID as the basis for designing their advanced control room for a next generation boiling water reactor plant. In addition to adopting the framework, they have also incorporated and adapted specific features of the EID interface for DURESS (e.g., the mass balance graphics) into some of their displays. This application is notable since it has been conducted at the scale of a full-scope nuclear power plant simulator. It is important to note, however, that this application of EID has yet to be evaluated.

More recently, Mitsubishi Atomic Power Industries in Japan has demonstrated a very strong interest in EID. They have contracted Battelle to initiate a 5 year research program, solely on EID. The program has just started, so no results are available yet.

In summary, EID has already attracted a fair amount of attention, at least from the nuclear industry. This is notable for several reasons. First, most of the research on EID has been conducted on a simplified yet representative process control simulation. This shows that it is possible to conduct laboratory research that has meaningful implications for complex, applied problems. Second, technology transfer from basic research to industry has been very hard to come by in human factors. The limited success of EID therefore stands out from the norm.

Nevertheless, it is important to point out that the true value of EID to industry has yet to be firmly established. It would certainly be very premature to suggest that all control rooms for nuclear power plants, for example, should be based on EID. Many research issues still remain to be addressed. More importantly, EID has yet to be evaluated on a large scale with professional operators. Thus, while there are many reasons to think that EID is a promising avenue to pursue, it is much too early to recommend it as a proven design framework.

5. FUTURE RESEARCH

Given the results of the work reviewed above, several issues for future research suggest themselves:

1. The biggest impediment to conducting research in this area is the lack of sensitive tools and methods for analyzing subject performance in dynamic, multi-degree of freedom systems. Any progress in this area is likely to yield enormous benefits.

2. It seems that people differ greatly in the extent to which they actively reflect upon, and learn from, their experience with an EID interface. Those who adopt a reflective attitude towards learning are the ones who are able to take the greatest advantage of an EID interface. It would be very useful to have a measure of this ability.

3. The introduction of EID interfaces seems to require a fundamental shift away from procedural thinking to adaptation and reflection on the part of

operators. Thus, it is likely that the benefits of EID interfaces cannot be fully realized with traditional training programs. The interaction between interface design and training needs to investigated.

4. The utility of EID principles needs to be investigated in broader, more realistic contexts. For example, the issue of how to integrate physical and functional information in larger-scale systems, where the entire system cannot be represented on a single display page, needs to be investigated further. Also, the utility of the EID framework for work domains other than process control needs to addressed.

Some of these issues are currently being addressed at the CEL at the University of Toronto.

ACKNOWLEDGEMENTS

This research was sponsored by research grants from ABB Corporate Research - Heidelberg, and the Natural Sciences and Engineering Research Council of Canada, and by research contracts with AECL Research and the Japan Atomic Energy Research Institute. The contributions of Klaus Christoffersen, Nick Dinadis, Chris Hunter, Bill Pawlak, and Alex Pereklita to this research are gratefully acknowledged.

REFERENCES

Bisantz, A. M., and Vicente, K. J. (1994). Making the abstraction hierarchy concrete. International Journal of Human-Computer Studies, **40**, 83-117.

Brunswik, E. (1956). Perception and the representative design of psychological experiments. (2nd ed.). Berkeley, CA: University of California Press.

Christoffersen, K., Hunter, C. N., and Vicente, K. J. (1994b). Research on factors influencing human cognitive behaviour (CEL 94-05). Toronto: University of Toronto, Cognitive Engineering Laboratory.

Dinadis, N., and Vicente, K. J. (1994). Ecological interfaces for complex industrial plants (CEL 94-06). Toronto: University of Toronto, Cognitive Engineering Laboratory.

Monta, K., Takizawa, Y., Hattori, Y., Hayashi, T., Sato, N., Itoh, J., Sakuma, A, and Yoshikawa, E. (1991). An intelligent man-machine system for BWR nuclear power plants. In Proceedings of AI 91: Frontiers in Innovative Computing for the Nuclear Industry (pp. 383-392). Jackson, WY: ANS.

Rasmussen, J. (1986). Information processing and human-machine interaction: An approach to cognitive engineering. New York: North Holland.

Vicente, K. J., and Rasmussen, J. (1990). The ecology of human-machine systems II: Mediating 'direct perception' in complex work domains. Ecological Psychology, **2**, 207-250.

Vicente, K. J., and Rasmussen, J. (1992). Ecological interface design: Theoretical foundations. IEEE Transactions on Systems, Man, and Cybernetics, **SMC-22**, 589-606.

Vicente, K. J. (1992a). Memory recall in a process control system: A measure of expertise and display effectiveness. Memory and Cognition, **20**, 356-373.

Vicente, K. J. (1992b). Multilevel interfaces for power plant control rooms I: An integrative review. Nuclear Safety, **33**, 381-397.

Vicente, K. J. (1992c). Multilevel interfaces for power plant control rooms II: A preliminary design space. Nuclear Safety, **33**, 543-548.

Vicente, K. J., Christoffersen, K., and Pereklita, A. (in press). Supporting operator problem solving through ecological interface design. IEEE Transactions on Systems, Man, and Cybernetics.

Woods, D. D. (1984). Visual momentum: A concept to improve the cognitive coupling of person and computer. International Journal of Man-Machine Studies, **21**, 229-244.

DEVELOPMENT OF ANALYSIS SUPPORT SYSTEM FOR MAN-MACHINE SYSTEM DESIGN INFORMATION

Hidekazu Yoshikawa, Makoto Takahashi
Institute of Atomic Energy, Kyoto University,
Gokasho, Uji,-shi, Kyoto-fu, 611 Japan
Kazunori Sasaki, Tohru Itoh, Takashi Nakagawa
Research Institute of Industrial Systems, Mitsubishi Electric Corporation,
Tukaguchi-Honmachi 8-1-1, Amagasaki-shi, Hyogo-ken, 661 Japan
Kazuhiro Kiyokawa, Akira Hasegawa
Institute of Human Factors, Nuclear Power Engineering Corporation,
Toranomon-3-17-1, Minato-ku, Tokyo, 105 Japan

Abstract: An integrated software system has been under development which aims at analyzing and evaluating the effectiveness of man-machine system designing, by computer simulations from varoius viewpoints of human factors. The target software system consists of two functional blocks: (i) distributed simulation system for man-machine interaction at man-machine interface, and (ii) man-machine design information evaluator. In this paper, the configuration of distributed simulation system is firstly introduced, followed by the explanation how operator simulator model is organized by petri net model.

Keywords: Man/machine systems, Man/machine interaction, Man/machine interface, Human factors, Human error, Cognitive systems

1. INTRODUCTION

Owing to the recent progress in computer control, information processing and human interface devices, designing of instrumentation and control (I&C) system for various plant system is rapidly going toward full digital I&C system with increased portion of automation. This is called as "supervisory control" (Sheridan,1992) in man-machine system (MMS), and the problems of I&C System designing in these days have centered on how to evaluate the man-machine interface (MMI) design from various human factors viewpoints, such as (i) appropriateness of operators' role in total man-machine system, (ii) evaluation of effectiveness for the operator's task fulfillment, (iii) evaluation of impact of human reliability by the introduction of new operation procedure.

The authors had made preparatory analysis to understand the human-machine interaction structure in the existing emergency operation procedure of nuclear power plant (Yoshikawa,1992). This was made by a graphical task analysis procedure which is comprised of (i) task transition diagram, (ii) hierarchical task analysis diagram, (iii) crew organization and communication diagram as input information for analysis, while (iv) action mode analysis table as output product. And then, from such desk-top analysis procedure which has been traditional in human factors analysis, the authors have proceeded to the development of computer-aided, simulation-based evaluation support system, for the predictive analysis of man-machine system designing from various viewpoint of human factors, especially centering on "cognitive mismatch in human-machine interaction". We are now in the process of developing the whole software system, the principal functional parts of which are divided into the two sub-systems, distributed simulation system and man-machine design information evaluator. The distributed simulation system is the back-end simulation system for all the three behaviors in concern in total man-machine system, ie., plant behavior, behavior of man-machine interface equipment , and operator's cognitive behavior. On the other hand, the man-machine design information evaluator is the front-end analysis support interface for both the qualitative and quantitative evaluations of cause and consequence of potential human error by operator's cognitive mismatch between perceived information and the presented information through man-machine interface.

In this paper, the methodology for organizing the distributed simulation system will be presented, with highlighting the modeling and simulation method of operator cognitive behavior at man-machine interface by petri net model

(Peterson, 1981).

2. DISTRIBUTED SIMULATION SYSTEM

Method of integrated dynamic simulation for man-machine system as a whole will be a straight forward and thus effective tool for the predictive analysis of the cognitive mismatch between the both behaviors of operator and plant system. On the part of plant system simulation, the related modeling methods and computer simulation techniques have been the area of traditional system engineering, and there are a lot of plant simulators available as software tools.

Concerning the behavior of human element, it is completely different from the modeling of machine element with respect to principle and mechanism of behavior, and it is a very difficult problem to model human cognitive behavior at man-machine interface. The authors have newly applied petri net for modeling human cognitive behavior, because of its easiness to organize the distributed simulation system in concern, considering the variety of human model level from a rather simpler one to more elaborate one in order to implement it into the whole system gradually. The authors are now under development of the initial phase of the operator simulator by petri net model which still takes into account of necessary psychological factors of operator cognitives such as perception, cognition and motor action with different mechanisms of human memory. The details of modeling framework, simulation method and the current status of development will be described separately in Section 3.

Apart from the problem of human modeling mentioned above, we will need a model of man-machine interface, when we think of integrated simulation on man-machine interaction. This should be an abstracted model of real man-machine interface equipment (ie. control board) which can effectively interface the model parameters in the both simulators of plant and operator. For this purpose, we have been developing "Man-Machine Interface (MMI) information generator", a kind of online knowlegde database model. By this MMI information generator, multifacet conditions of man-machine interface equipment with respect to structural configuration, topological relationship, functional characteristics, etc., are modeled by a knowledge database with hierarchical frame representation, and the dynamic information elements coming from and going to the both simulators can be included in the appropriate slots of the frames and can be updated from time to time during dynamic simulation.

To sum up, the computer simulation of man-machine interaction will be organized by the distributed architecture of the three simulators; plant simulator, operator simulator and MMI information generator. Figure1 summarizes general idea of this distributed simulation system with the information flow among the three different simulators.

3. OPERATOR SIMULATOR

3.1 Conceptual framework of human model by petri net

Reason who viewed human as "fallible machine" proposed a general framework of human modeling which will be able to predict not only right human performance but also possible human error form (Reason, 1990). This is the product of conceptual aggregation from the existing knowledge in the field of cognitive psychology, and from his thinking, the

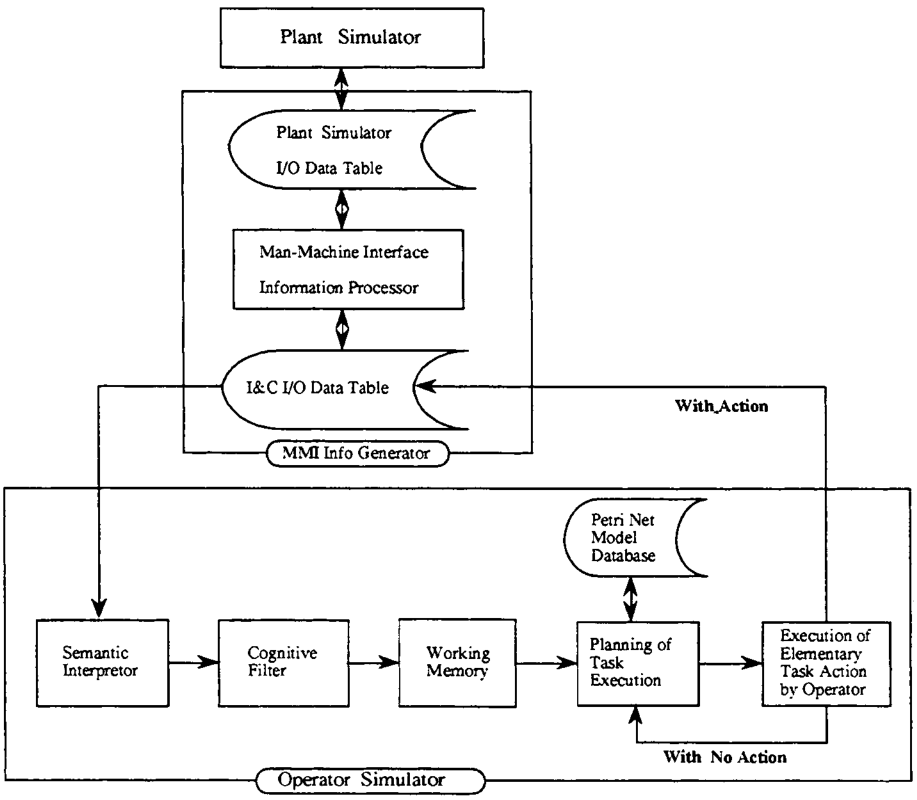


Fig. 1. Organization and information flow in distributed simulation system.

general idea on how to model human cognitives at man-machine interface can be summarized as shown in Fig. 2.

There have been a lot of simulation works published thus far for modeling human cognitive behavior, and most of them applies AI technique using blackboard architecture. But the authors set to develop operator simulator model by applying petri net model because of the following considerations to implement the general framework in Fig.2 into an efficient distributed simulation system:

(1) Need of modeling both serial processing portion (conscious process) and paralell processing (unconscious process),
(2) Need of describing "chunking of knowledge structure" by hierarchical way,
(3) Need of visualizing state transition in cognitive process,
(4) Ease of modeling the interaction of cognitive process with out-world information through MMI information generator,
(5) Ease of handling knowlegde base structure (update and grade-up) to meet the analysis objectives flexibly, and
(6) Need of estimating various performance measures (such as mental worklord index).

The advantage of applying petri net model lies in its merit of describing state transition by the use of "place" and "transition", where we can mix both the serial and parallel processes, and visualize the structure and the dynamic process of state transition comprehensively on graphical display.

But there remains problems if we proceed to apply ordinary petri net model for formulating the conceptual model as described in Fig.2. The most important point is that the distinction between "processing mechanism " and "data structure " is not necessarily clear from Fig.2 because it is rather crude picture on how human cognitives work as a whole. Therefore, we have to make clear the general idea in Fig. 2 into a more workable formulation to construct human model by petri net model.

At the moment, the authors are thinking of translating it into the system architecture as shown in Fig.3, where the whole functions of human cognitives are divided into perception system, cognitive system and motor system. In Fig.3, the distinction between processing mechanism and data structure is only made for cognitive process by "KB processing" and "KB database", respectively.

The area of PWM is the background pallete area of information elements coming from perception system, to and from KB database and going to motor system, while FWM that for conscious processing to the attended information elements from PWM. KB database corresponds to long-term memory.

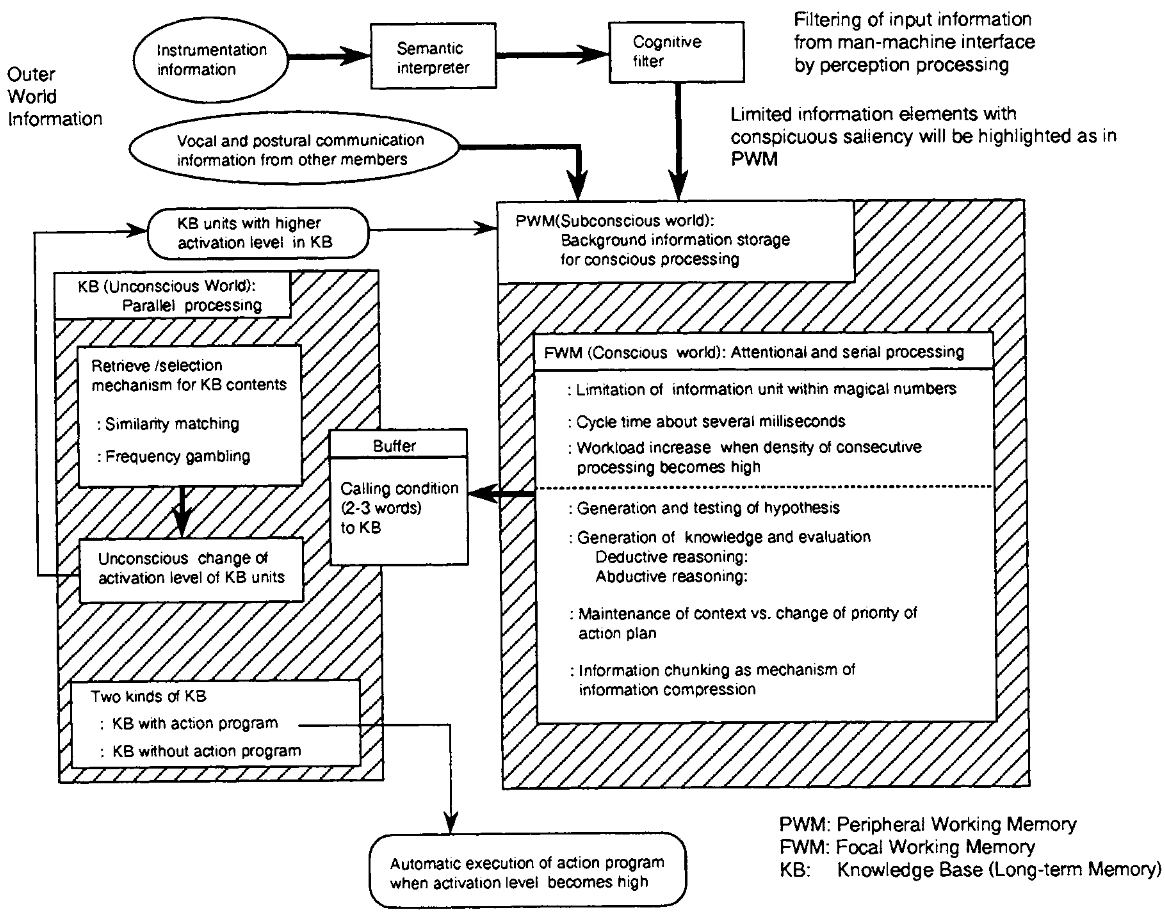


Fig.2. Conceptual framework of human model for man-machine interface problem.

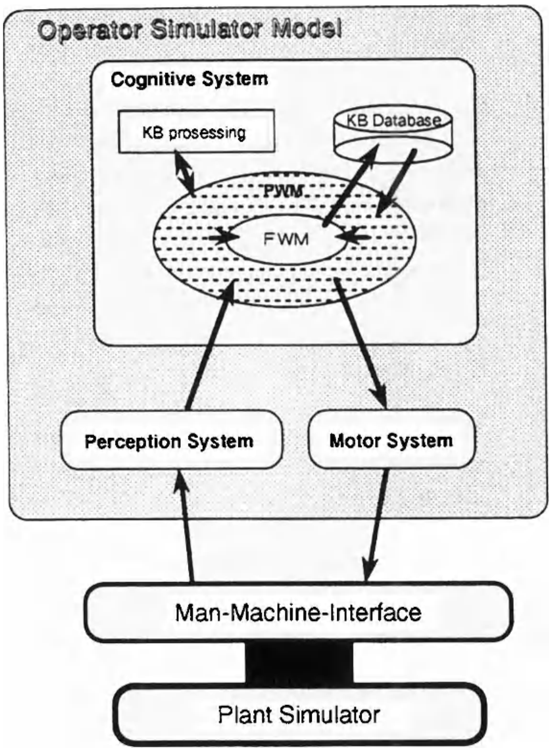


Fig.3. Configuration of operator simulator model.

Items	Contents
Information Content	Which does what?
Importance Index	Priority measure of processing
Information Source	Perception System, Knowledge-Base, Knowledge Prosessing
Status of KB Processing	Prosessing is needed, Processing is ended, Processing is unknown
Relationship between Other Information	

Fig.4. Information elements for FWM pallete.

The important functions in KB processing are (i)prioritization of incoming information elements which come from perception system and KB database, by assigning "Importance Index". (The information elements with high priorities are those with "High Saliency" from perception system all the way through semantic interpreter and cognitive filter, and spontaneously activated knowledge elements with "High Activation Level" in KB database. The information elements or words with high "Importance Index " correspond to "Calling Conditions" in Fig.2.), (ii)retrieval of information elements from KB database by keywords search with Calling Condition, (iii)inference function at FWM with interaction to KB database keeping the context in concern, (iv)chunking function of the both information elements activated in FWM during attentional process and the knowlege elements in KB database which are "theorized " unconsciously as learning or experience will accumulate. If the cognitive process bypasses FWM and the action is made only through PWM pallete, then this process will be "skill-based process" and if the process is on FWM but it follows just the procedural knowledge base in KB database with no further inference mechanism triggered, then it will be "rule-based process" otherwise "knowledge-base".

3.2 Petri net model for cognitive system

We have conducted on the petri net modeling for the part of cognitive system with rather simpler modeling assumptions than those required for the final targets as mentioned above; only rule-based procedural process with a single calling condition (single word) selected by preset "Importance Index" vector. The items and contents of the information elements in FWM are shown in Fig.4, while the way how to convert a given operation procedure into petri net model as information elements in KB database is illustrated as shown in Fig. 5.

As seen in Fig.5, there are two types of places, simple places indicated by single circles,and hierarchical places by double circles. Note that a statement label attached to each place is the statement of specific action to be made. The content of hierarchical place will be further expanded like the sequence of places and transitions as shown in the bottom part of Fig. 5. The transitions are indicated by perpendicular bars as shown in the same figure.

There are four types of tokens for place, as explained in Fig.6. If all the places connected to a transition towards upstream direction become "finished" state, then the transition will fire and all the places downwards becomes "candidate" state. The state of place changes from "candidate" to "finish", if the content of attended information element (highest Importance Index in FWM list) agrees with the statement label among the places with " candidate " state.

The way how PWM, FWM and KB database work together by the cognitive system is illustrated in Fig.7. The sequence of transient is as follows;

(1)In KB database, "instrument A low alarm" is already at "candidate" state. First, an information element "Instrument A Alarm on" is noticed via perception system and PWM, and this changes highest "Importance Index" vector in FWM so that the most attended action in FWM is "Instrument A alarm NEEDED".

(2)Then FWM starts information search to KB database by keyword "Instrument A alarm".

(3)The search in KB database finds that the place "Instrument A Low Alarm"with "candidate " state (mark α) agrees with the keyword. Then this place α will be activated.

(4)The token of the place α changes to "finished" state in the meantime and the downstream transition fires. Then all the places downstream of the transitions (β and γ) become "candidate" state. Afterwards, those status change information in KB database are sent back to FWM via PWM, so that the FWM information content be altered as "Instrument A alarm ENDED", "Assume Accident-type A NEEDED", and "Assume False Alarm NEEDED".

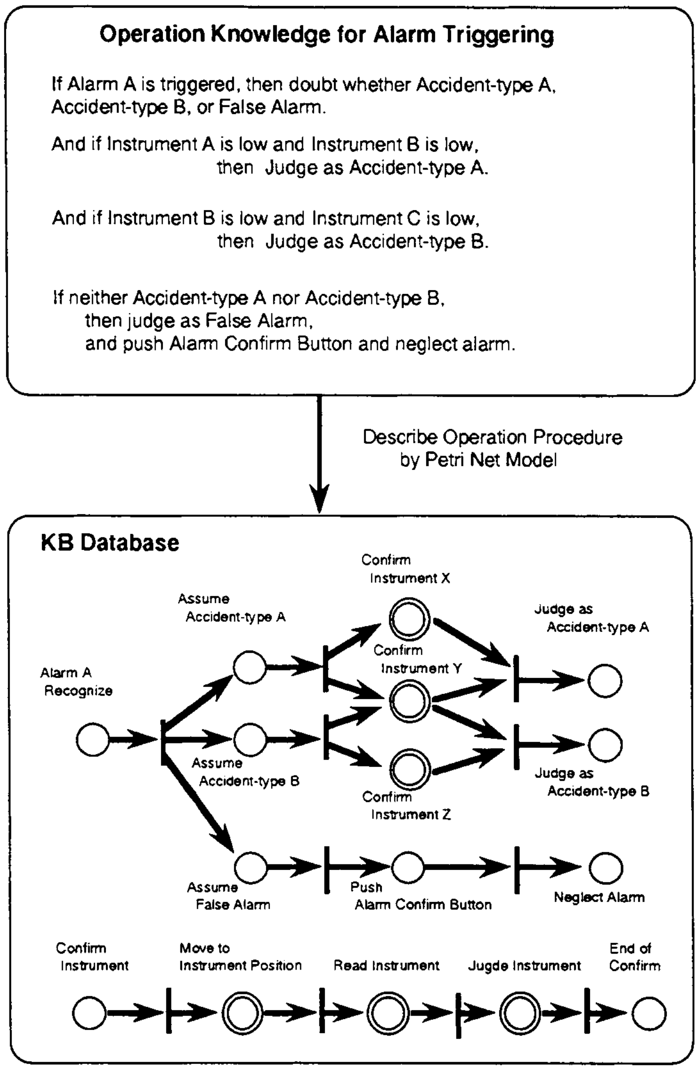


Fig.5. Method of converting operation procedure to Petri Net model as KB database.

Type of Token	Status of Place	Meaning
○	"Candidate" of Action	If the place content which is already "candidate" state agrees with the request from FWM, then the token of this place wil become "Finish" state, after action succeeds. But if the action failed, the token becomes "Failed" state. As soon as the upstream transition is fired, the place becomes "Candidate" state and the information content of the place is sent to PWM. (If the place is the initial one of the petri net sequence with no token marks in the sequence, it is assumed as "candidate" status .)
●	Action "Finished"	When the place becomes this state, then the direct downstream transition to this place is checked whether or not it satisfies firing condition, and if it fires, then the output places of this transition become "Candidate" state. As soon as the place is found to shift from "Candidate" to "Finished" status by the request from FWM, status flug of KB processing in FWM is requested to change from "needed" to "ended" by way of PWM.
×	Action "Failed"	When the place becomes this state, the request of changing the status flug of KB processing in FWM from "needed" to "ended" is made via PWM. No firing of the downstream transition occurs in this case.
None	No activating state	The place with this token state is outside of the search from FWM.

Fig.6. Types of tokens and their meaning.

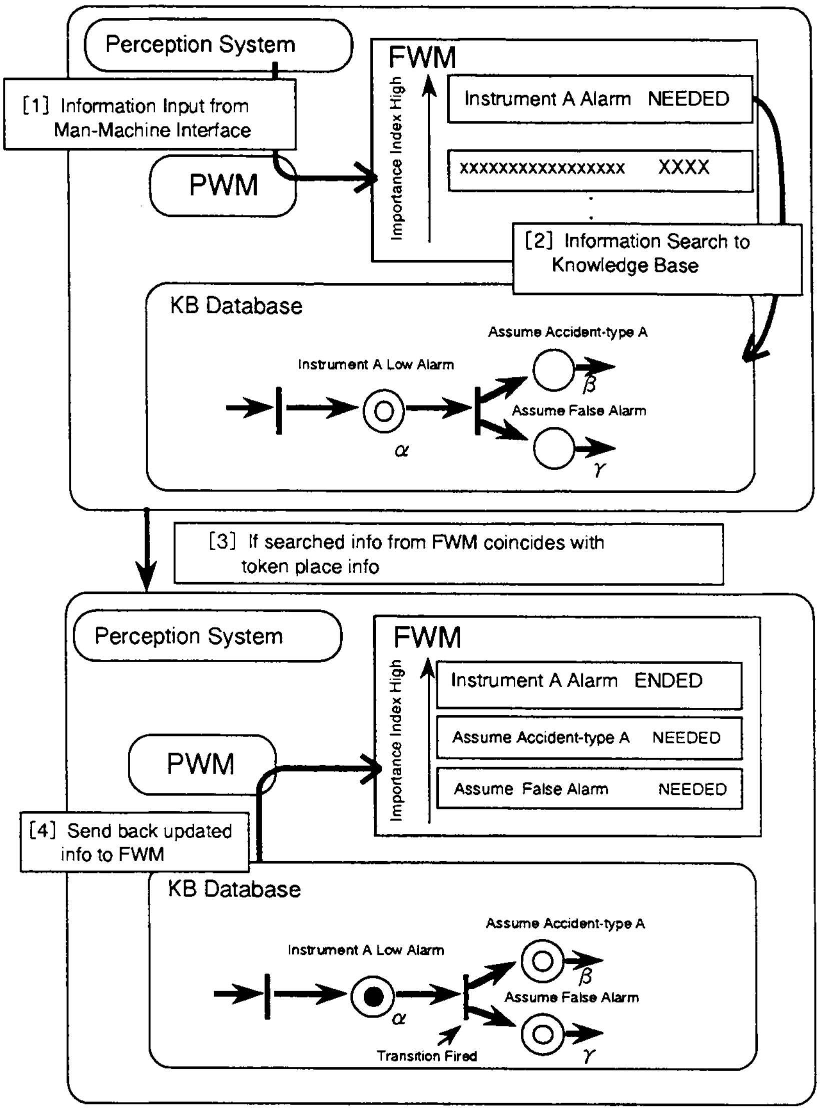


Fig.7. Control algorithm on how PWM, FWM and KB database work together in cognitive system.

4. CONCLUSION

In this paper, the authors presented the modeling methods and development of the distributed simulation system for man-machine interaction, especially highlighting the use of petri net model for organizing cognitive system in the operator simulator. Although the software system development is rather initiating phase, the example simulation practise thus far conducted for PWR LOCA accident resulted in that the proposed method would be fundamentally an effective approach alternative to the AI methodology to model various aspects and characteristics of human cognitive behavior at man-machine interface.

In future, the authors are going further development such as elaboration of the modeling capabilities of cognitive system, development of the remaining sub-system elements such as perception system and motor system, and finally total system integration, in order to realize the target distributed simulation system of man-machine interaction as the back-end system of the whole analysis support system for man-machine system design information.

REFERENCES

Peterson, J.L.(1981). *Petri Net Theory and the Modeling of Systems*, Prentice-Hall, New York.

Reason, J.(1990). *Human Error*, Cambridge University Press, Cambridge.

Sheridan, T.B. (1992). *Telerobotics, Automation, and Human Supervisory Control*, The MIT Press, Cambridge.

Yoshikawa, H., Gofuku, A., Itoh, T. and Sasaki, K. (1992). An Investigative Study Towards Constructing Anthropocentric Man-Machine System Design Evaluation Methodology. *Proc. Post-ANP'92 Conference Seminar on Human Cognitive and Cooperative Activities in Advance Technological Systems*, (Kondo,S. (Ed.)) pp. 25-36, Atomic Energy Society of Japan, Tokyo.

A DESIGN METHOD FOR INCORPORATING HUMAN JUDGMENT INTO MONITORING SYSTEMS

K. TANAKA*, and G.J. KLIR**

**Ibaraki University, Department of Computer and Information Sciences, Hitachi, Ibaraki 316, Japan*

***State University of New York, Department of Systems Science and Industrial Engineering, Thomas J. Watson School of Engineering and Applied Science, Binghamton, NY 13902, USA*

Abstract. In safety monitoring, uncertainty domain exists in states of the monitored object where a sensor cannot determine whether the object is in safety or in danger. To reduce the uncertainty, a new design method of non-homogeneous monitoring system(NHMS) is proposed that utilizes two types of sensors which have different thresholds. Our paper analyzes how to judge states of the object at uncertainty domain in NHMS without information added from human being and in NHMS with it, respectively. Probabilistic model and Dempster-Shafer model are adopted and the conditions are proposed under which inspections by human beings are effective.

Key Words. Man/machine systems; Probabilistic models; Decision theory; Safety analysis; Monitoring systems; Dempster-Shafer theory; Uncertainty

1. INTRODUCTION

Reliable safety monitoring is important for detecting abnormal situations in large-scale systems or in those systems that cannot be observed directly. It is known, however, that a sensor is vulnerable to two types of contradictory failures; a failed dangerous (FD) failure and a failed safe (FS) failure (Henley, 1985). An FD failure means that a sensor fails to generate an alarm signal for the monitored object in danger, while an FS failure means that a sensor generates an alarm signal for the object in safety. A safety monitoring system, which has been developed to prevent both FD and FS failures of individual sensors, generates a reliable alarm through a synthetic judgment based on information from several sensors.

Phillips (1980) has shown that the optimal structure of N independent identical sensors is a k-out-of-N, under which an alarm is generated if at least k sensors produce alarm signals. Though this structure is effective to avoid influence from physical noise or transient faults, FD/FS failures sometimes occur even if a sensor doesn't fail, because it is difficult in some states of the monitored object for a sensor to judge whether the object is safe or dangerous. This paper focuses on this **uncertainty zone**. To reduce ambiguity in the uncertainty zone, Tanaka (1994) proposed a reliable design method for monitoring systems that consist of two different types of sensors. Such systems are called in this paper as **non-homogeneous monitoring systems(NHMS)**. NHMS have two properties; one is to decrease the occurrences of conflict outputs of a sensor and the other is to easily identify the uncertainty zone. The latter property provides a chance to utilize information obtained by human judgments based upon detailed inspections of the monitored object.

Our paper describes an optimal design method of NHMS which minimizes the expected value of losses from FD/FS failures. The basic issue is to decide when the system alarm should be generated on the basis of monitored states. After discussing the design of machine system, this paper analyzes the design of human-machine system via two approaches, one based on probability theory and one on the Dempster-Shafer theory. Conditions are shown under which inspection by humans is effective in NHMS.

2. NON-HOMOGENEOUS MONITORING SYSTEMS

2.1. *SP sensor and FW sensor*

Two-state models do not consider uncertainty zone in which neither safety nor danger is affirmed by sensors since the object is required to be always either safe or dangerous in their models. To consider FD/FS failures derived from uncertainty zone, lack of information about danger must be distinguished from safety since not to detect dangerous phenomenon of the object doesn't always guarantee its safety.

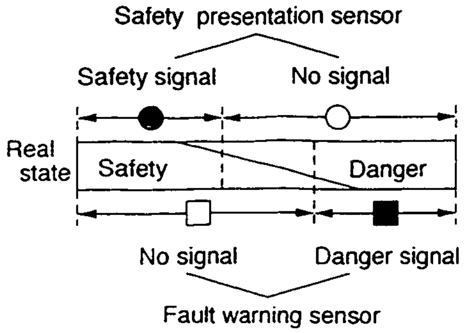


Fig. 1 SP sensor and FW sensor

To deal with this issue, Tanaka (1994) introduced two kinds of sensors: a safety-presentation (SP) sensor and a fault-warning (FW) sensor. An SP sensor generates a safety signal(say S) only when it confirms that the monitored object is safe. This means that even if the SP sensor doesn't generate a safety signal($\sim S$), it doesn't always reveal danger. On the other hand, an FW sensor generates a danger signal(D) only when it confirms that the object is dangerous. In the FW sensor, no alarm signal($\sim D$) doesn't always guarantee safety. It is expected that few FS failures occur in the SP sensor and that few FD failures occur in the FW sensor. Accordingly, an NHMS that consists of both types of sensors is expected to decrease the occurrences of both FD failures and FS failures in the uncertainty zone. An NHMS is easily obtained by setting different thresholds of sensors(Fig.1).

Consider two-sensor system which has three types of combination: two SP sensors, two FW sensors, and an NHMS. Our first problem is to decide whether or not an NHMS is optimal among them. That is the **optimal combination problem (OCP)** of deciding the optimal combination to minimize damage from FD/FS failures among three combinations. The OCP is dealt with in terms of the Bayesian methodology in Sec. 3.

2.2. *Inspection activity by human*

If an NHMS is a machine system which automatically generates a warning for danger based on sensor outputs, it may be unable to utilize information about uncertainty zone because it interprets the whole uncertainty zone as either safe state or dangerous one. Most valuable method in such cases is to incorporate human judgments based on detailed inspection of the monitored objects.

Fig.2 shows four output patterns of an NHMS. Pattern (a) is a standard pattern for the safe object since the SP sensor generates a safety signal and the FW sensor doesn't generate a danger signal($S\cap \sim D$). Similarly, pattern (b) expresses a standard pattern for the dangerous

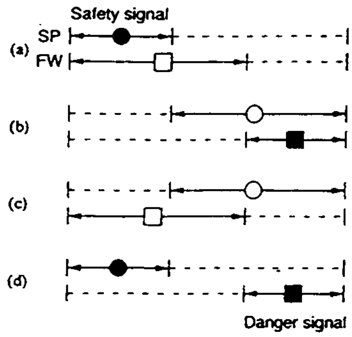


Fig. 2 Output patterns of NHMS

object($\sim S \cap D$). Pattern (c) reveals that neither the SP sensor confirms safety of the object nor the FW sensor confirms the existence of a dangerous phenomenon($\sim S \cap \sim D$). This reveals that the state is in the uncertainty zone and that it is difficult to judge correctly whether the object is safe or dangerous. In this case, if there is enough time to inspect the object and the inspection is performed with low cost and low probability of human error, it is desired to judge from added information after detailed inspections. Our interest is to know the conditions under which inspection is better from the point of balance of cost and performance. In pattern (d), it seems impossible to judge whether the object is safe or dangerous($S \cap D$). This pattern presents information about physical failures in sensors, which should be repaired of physical trouble immediately.

Thus, (c) and (d) reveal uncertainty zone and sensor failures, respectively. After a Bayesian approach to analyzing these patterns, based on Tanaka (1994), is discussed in Sec.3, the use of the Dempster-Shafer theory is examined in Sec.4.

3. BAYESIAN PROBABILISTIC APPROACH

3.1. *Probabilistic Model*

Consider a general safety monitoring system consisting of n sensors which monitor the same characteristic. X is a set of states of the monitored object in the real world. Assume that X is divided into safety domain S and danger domain D, that is, $X = S \cup D$ and $S \cap D = \phi$. Each sensor monitors a totally ordered value $\eta(x)$ for the real state $x \in X$. Fig.3 shows an NHMS model which consists of two sensors.

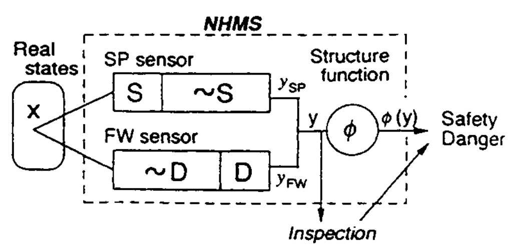


Fig. 3 Non-homogeneous monitoring system model

$y_{SP}(i)$	: output variable of the i-th SP sensor
$y_{FW}(j)$	: output variable of the j-th FW sensor
T_{SP}, T_{FW}	: thresholds of SP/FW sensors
$y_{SP} = 0$	: safety (S), when $\eta(x) \leq T_{SP}$
$y_{SP} = 1$	: not safety ($\sim S$), when $\eta(x) > T_{SP}$
$y_{FW} = 0$	: not danger ($\sim D$), when $\eta(x) \leq T_{FW}$
$y_{FW} = 1$	: danger (D), when $\eta(x) > T_{FW}$

Let $\mathbf{y}$ be a vector of output variables of sensors : $\mathbf{y} = (y_{SP}(1), .., y_{SP}(k), y_{FW}(k+1), .., y_{FW}(n))$, and ϕ be a structure function: $\mathbf{y} \rightarrow \{0, 1\}$.

$\phi(\mathbf{y}) = 0$	: no system alarm generation
$\phi(\mathbf{y}) = 1$	: system alarm generation
C_a	: loss caused by an FS failure
C_b	: loss caused by an FD failure
π	$:= p(x \in D)$.

The expected loss I of ϕ is expressed as

$$I = K \sum_{\mathbf{y}} \overline{\phi(\mathbf{y})} p(\mathbf{y}|_D) + L \sum_{\mathbf{y}} \phi(\mathbf{y}) p(\mathbf{y}|_S), \quad (1)$$

where $\overline{\phi} = 1 - \phi$, $K = \pi C_b$, $L = (1 - \pi) C_a$, and $p(\mathbf{y}|_D) = p(\mathbf{y}|x \in D)$. The two terms in (1) express the expected losses causes by FD failures and FS failures. The optimal system is required to minimize I. In terms of decision theory, the optimal system is the Bayesian risk that minimizes the Bayesian expected risk I. The optimal structure function ϕ which attains Bayesian risk is easily obtained by the following rule(Henley, 1985):

$$\phi(\mathbf{y}) = \begin{cases} 1 & \text{if } g(\mathbf{y}) > 0 \\ 0 & \text{if } g(\mathbf{y}) \leq 0, \end{cases} \quad (2)$$

where $g(\mathbf{y}) = K \cdot p(\mathbf{y}|_D) - L \cdot p(\mathbf{y}|_S)$, which is derived from the rewritten form of (1),

$$I = K - \sum_{\mathbf{y}} \phi(\mathbf{y}) g(\mathbf{y}). \quad (3)$$

This rule is based on a function of all sensor outputs which include contradicted patterns.

OCP is to minimize I by controlling the combination of SP sensors and FW sensors, which means controlling values of $p(\mathbf{y}|_S)$ and $p(\mathbf{y}|_D)$. Assume that all sensors have the same distribution. Let the conditional distribution of $p(t|_S)$ for observed value $t = \eta(x)$ be $F(t)$ and that of $p(t|_D)$ be $G(t)$.

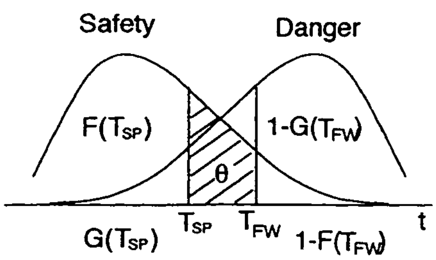


Fig. 4 Distribution functions of sensor outputs

Since $F(T_{SP}) = p(\eta(x) < T_{SP}|_S)$,

$$\begin{array}{lll} p(y_{SP} = 0|_S) & = F(T_{SP}) & (= a) \\ p(y_{FW} = 0|_S) & = F(T_{FW}) & \\ p(y_{SP} = 0|_D) & = G(T_{SP}) & (= c) \\ p(y_{FW} = 0|_D) & = G(T_{FW}). & \end{array} \quad (4)$$

3.2. *Machine Monitoring System*

Consider the OCP for a basic two-sensor system of selecting the optimal type among three types of SS, DD, and NHMS, given distributions F, G and thresholds T_{SP}, T_{FW}. Now, assume that

(A1) sensors are mutually independent,
(A2) $0 < \theta < 1$, where $F(T_{FW}) - F(T_{SP}) = \theta$,
(A3) $F(T_{SP}) \geq 1 - F(T_{FW})$.

In (A2), the value θ shows probability of the uncertainty zone. If $\theta = 0$, an SP sensor would be not distinguishable from an FW sensor.

Moreover, to avoid needless complexity in our model, assume symmetry for two the distributions, F and G, with respect to the thresholds;

$$(A4) \begin{cases} F(T_{SP}) = 1 - G(T_{FW}) & = a, \\ G(T_{SP}) = 1 - F(T_{FW}) & = c. \end{cases}$$

Then, $G(T_{FW}) - G(T_{SP}) = \theta$ also follows from (A2) and (A4). By introducing two parameters,

$$\gamma = \frac{F(T_{SP})}{G(T_{SP})} = \frac{a}{c}, \quad \delta = \frac{K}{L} = \frac{\pi C_b}{(1 - \pi) C_a}, \quad (5)$$

the following Proposition 3.1 provides us with a criteria for selecting the optimal combination in the case of $\delta \geq 1$(see Tanaka (1994) for $\delta < 1$), where $A(\delta) = \delta^2 - (\delta - 1)(1 - \sqrt{1 + \delta^2})$, $B(\gamma, \delta) = \frac{\delta - \gamma}{1 - \gamma\delta}$, $C(\gamma, \delta) = \frac{2\gamma(\delta - 1)}{(\gamma - 1)(\gamma + \delta)}$, and $D(\gamma) = \frac{\sqrt{2}}{\sqrt{\gamma} + \sqrt{\gamma^{-1}} + \sqrt{2}}$.

Proposition 3.1 The optimal combination system and the optimal structure for a two-sensor system in $\delta \geq 1$ are specified in the following figure:

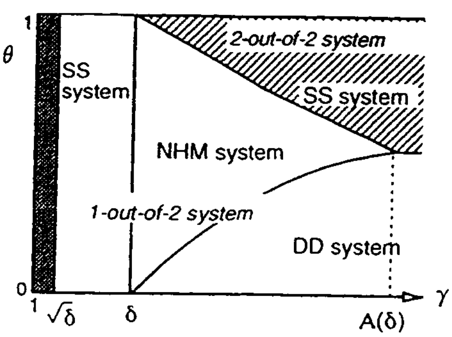


Fig. 5 The optimal combination system

Proposition 3.1 reveals that NHMS is optimal when $\delta \leq \gamma \leq A(\delta)$ and $B(\gamma, \delta) \leq \theta \leq C(\gamma, \delta)$.

The comparison is meaningful only when $\gamma^{-2} < \delta \leq \gamma^2$, because when $\gamma^{-2} \geq \delta$, no systems generate a system alarm and when $\gamma^2 < \delta$, all the systems always generate a system alarm. An interesting result is that the optimal combination is uniquely determined regardless of the value of uncertainty zone θ when $\gamma \in [\sqrt{\delta}, \delta]$.

Example 1 Suppose $\pi = 0.001, C_a = 1$, and $C_b = 20,000$ which leads to $\delta = 20.0$. When $(F(T_{SP}), \theta, G(T_{SP})) = (0.85, 0.1, 0.05)$, SS system is optimal independent of θ since $\gamma = \frac{0.85}{0.05} = 17 \in [4.47, 20.0](= [\sqrt{\delta}, \delta])$. When $(F(T_{SP}), \theta, G(T_{SP})) = (0.89, 0.1, 0.01)$, NHMS is optimal since $\gamma = 89 \in [20.0, 761.5(= A(20))]$ and $B(89, 20) = 0.039 < \theta < C(89, 20) = 0.353$

It should be remarked that the expression of the structure in SS is different from that in DD. For example, 2-out-of-2 system in SS means that when neither of the two SP sensors produces a safety signal, the system generates a system alarm.

Both the optimal combination and the optimal structure depend on three parameters, γ, θ, and δ, where γ is not an absolute value but a relative parameter(see (5)). Moreover, if it is assumed that

(A5) $G(x)F(x)^{-1}$ is strictly increasing,

one parameter is eliminated.

Proposition 3.2 Under fixed F and G, determining θ is equivalent to determining γ.

Prop. 3.1 and 3.2 lead to the result that the optimal combination depends only on the uncertainty zone θ and the environment parameter δ under fixed F and G.

3.3. *Human-machine Monitoring System*

As mentioned earlier, when neither the SP sensor nor the FW sensor generates a signal in the NHMS, the object is considered to be in the uncertainty zone. In this section, conditions are discussed under which a detailed inspection is effective to judge the state of the object.

Let C_I denote an inspection cost and R denote the probability either that the inspection cannot detect any actual danger or that damage will occur before inspection is complete. When inspection is performed, the inspection cost C_I is required even if the object is safe. Moreover, if the object is dangerous, the expected value of the damage from incomplete inspection RC_b is added to the inspection cost. If inspection is always complete, R should be set to zero. When we denote by I_{insp} the expected loss for the NHMS with inspection activities and by I_{NH} the expected loss without inspection, we have

$$I_{NH} - I_{insp} = \begin{cases} (L - RK - C_I)Z & \text{if } \delta \geq 1 \\ (K - RK - C_I)Z & \text{if } \delta < 1, \end{cases} \quad (6)$$

where $Z = F(T_{SP})\overline{F(T_{FW})} + \overline{F(T_{SP})}F(T_{FW})$. The right term in (6) reveals the effect of inspection. Since $Z \geq 0$, the next proposition is obtained. Note that the condition is independent of distributions F and G.

Proposition 3.3 NHMS with inspection is more desirable than NHMS without inspection if and only if $min(K, L) - RK > C_I$.

Example 2 Consider again Example 1. Suppose $R = 0.01$ and $C_I = 0.1$. Since $min(K, L) - RK = L - RK = 0.799 > C_I$, NHMS with inspection is better than that without it. When $(F(T_{SP}), \theta, G(T_{SP})) = (0.89, 0.1, 0.01)$, the effect by inspection is $I_{NH} - I_{insp} = 0.141 - 0.059 = 0.082$. Moreover, when R and C_I satisfy $0.999 > 20R + C_I$, inspection activities are desirable.

4. DEMPSTER-SHAFER THEORETICAL APPROACH

In the probabilistic model, machine system considers the whole uncertainty zone as a safe state or as a danger state and only human-machine system can divide it into safety part and danger part. In this section, Dempster-Shafer theory (D-S theory) is used to perform a division of the uncertainty zone by machine system. D-S theory has been utilized for uncertainty analysis (Klir, 1988). Throughout the whole section, (A1),(A2),(A3), and (A4) are assumed.

4.1. *Dempster-Shafer Theory*

Let

$$m : \wp(X) \to [0, 1] \quad (7)$$

be a function, referred to as a *basic probability assignment*(BPA), such that $m(\phi) = 0$ and $\sum_{A \in \wp(X)} m(A) = 1$. Note that the domain of BPA, $\wp(X)$, is different from the domain of a probability distribution function, which is X. One of the central issues in D-S theory is how to combine BPA's given by several independent sources. Suppose there are two BPA's, m_1 and m_2. The standard way of combining m_1 and m_2 is expressed by

$$q(C) = \sum_{C = A \cap B} m_1(A)m_2(B). \quad (8)$$

However, the combined belief must be normalized because $q(\phi) \geq 0$ in general. Inagaki (1991) proposed a combination rule which unifies Dempster's rule and Yager's rule. This paper uses another unified rule:

$$\begin{aligned} m_k(C) &= kq(C) \qquad \text{for } C \neq X \\ m_k(X) &= kq(X) + (1 - k(1 - q(\phi))), \end{aligned} \tag{9}$$

where $1 \leq k \leq \frac{1-q(X)}{1-q(X)-q(\phi)} \equiv k_{max}$. This unified combination rule, which is equivalent to Inagaki's rule, coincides with Yager's rule when $k = 1$ and with Dempster's rule when $k = (1 - q(\phi))^{-1}$; the rule for k_{max} is called an Extra rule (Inagaki, 1991). Yager's rule(Yager, 1987),

$$\begin{aligned} m(C) &= q(C) \qquad \text{for } C \neq X \\ m(X) &= q(X) + q(\phi), \end{aligned} \tag{10}$$

regards contradictions as coming from ignorance. Dempster's rule(Shafer, 1976),

$$m(C) = \frac{q(C)}{1 - q(\phi)} \quad \text{for all } C \in \wp(X). \tag{11}$$

constructs a combined BPA from that portion of information which does not exhibit any contradictions. The Extra rule is equivalent to adopting Dempster's rule for all sets except X so that the rule keeps only $q(X)$, i.e., $m_{k_{max}}(X) = q(X)$.

Belief and plausibility measures are uniquely determined for each $A \in \wp(X)$ by the formulas

$$\begin{aligned} Bel^k(A) &= \textstyle\sum_{B \subseteq A} m_k(B), \\ Pl^k(A) &= \textstyle\sum_{B \cap A \neq \phi} m_k(B) \end{aligned} \tag{12}$$

4.2. *Expected Loss in D-S Model*

D-S theory is suitable for modeling our problem because states of SP sensor are not expressed by S and $\sim S$ but by S and unknown X. Let us introduce some BPA's given S or D;

$$\begin{aligned} & m_{SP}(S|_S) = p(t \leq T_{SP}|_S) && (= a) \\ & m_{SP}(S|_D) = p(t \leq T_{SP}|_D) && (= c) \\ & m_{FW}(D|*) = p(T_{FW} \leq t|*) && (* = S \text{ or } D) \\ & m_{SP}(X|*) = 1 - m_{SP}(S|*) && \\ & m_{FW}(X|*) = 1 - m_{FW}(D|*) && \end{aligned} \tag{13}$$

Our first problem is to decide how to synthesize outputs of an SP sensor and an FW sensor(Fig.6),

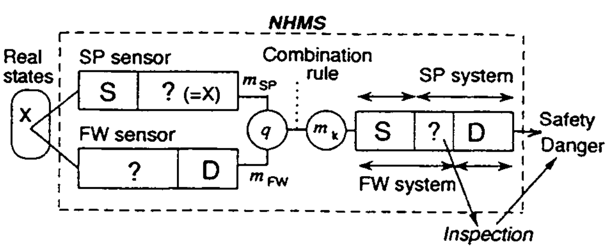


Fig.6 D-S model for NHMS

that is, to decide the optimal value of k in the combination rule(9) after $q_S(\cdot)$ is obtained from $m_{SP}(*|_S)$ and $m_{FW}(*|_S)$, and $q_D(\cdot)$ is obtained from $m_{SP}(*|_D)$ and $m_{FW}(*|_D)$ by (8). In this process, ambiguities in conflicting outputs and the uncertainty zone are partly deleted by interpreting them as a safety part or a danger part. Furthermore, according to (12), belief and plausibility measures given safe object are defined from the unified assignment by

$$\begin{aligned} Bel_S^k(D) &= m_k(D|_S) \\ Pl_S^k(D) &= m_k(D|_S) + m_k(X|_S). \end{aligned} \tag{14}$$

$Bel_D^k(S)$ and $Pl_D^k(S)$ are also defined similarly.

The expected loss in D-S model, J, is easily expressed by these measures: for the SP system type, it is

$$J_{SP}(k) = \pi C_b Bel_D^k(S) + \overline{\pi} C_a Pl_S^k(D); \tag{15}$$

for the FW system type, it is

$$J_{FW}(k) = \pi C_b Pl_D^k(S) + \overline{\pi} C_a Bel_S^k(D). \tag{16}$$

Our second problems is to decide the optimal type among SP system type and FW system type. In other words, it is to decide which the remaining part in the uncertainty zone, X, should be interpreted as safety or as danger.

4.3. *Machine Monitoring System*

The optimal value k in the unified combination rule(9) is obtained as follows.

Proposition 4.1
(1) For the SP system type,

$$min_k J_{SP}(k) = \begin{cases} J_{SP}(1) & \text{if } \delta \geq \frac{q_S(S)}{q_D(S)} \\ J_{SP}(k_{max}) & \text{if } \delta < \frac{q_S(S)}{q_D(S)} \end{cases}$$

(2) For the FW system type,

$$min_k J_{FW}(k) = \begin{cases} J_{FW}(k_{max}) & \text{if } \delta > \frac{q_S(D)}{q_D(D)} \\ J_{FW}(1) & \text{if } \delta \leq \frac{q_S(D)}{q_D(D)} \end{cases}$$

Since $q_S(S) = q_D(D) = a\overline{c}$, $q_D(S) = q_S(D) = \overline{a}c$, and $q_S(X) = q_D(X) = \overline{a}\overline{c}$ from the assumption of symmetry (A4), the optimal combination rule and the optimal system type are simply expressed by the following two propositions, where $\alpha = \frac{F(T_{FW})}{G(T_{FW})} = \frac{1-c}{1-a}$.

Proposition 4.2 Optimal combination rule is

(1) Yager's rule ($k = 1$) if $\delta \leq (\alpha\gamma)^{-1}$ or $\alpha\gamma \leq \delta$,
(2) Extra rule ($k = k_{max}$) if $(\alpha\gamma)^{-1} \leq \delta \leq \alpha\gamma$.

Proposition 4.3 Optimal system type is

(1) SP system type if $\delta \geq 1$
(2) FW system type if $\delta \leq 1$

Prop.4.3 reveals that the optimal type depends only on the ratio of K and L, and is not dependent on the assignment m.

Example 3 Consider again Example 1. Since $\delta = 20$, SP system type is optimal by Prop.4.3. Suppose that $m_{SP}(S|_S) = m_{FW}(D|_D) = 0.89$, $m_{SP}(X|_S) = m_{FW}(X|_D) = 0.11$, $m_{SP}(S|_D) = m_{FW}(D|_S) = 0.01$. Then, Extra rule is optimal by Prop.4.2 since $\alpha\gamma = 801$ and $(\alpha\gamma)^{-1} = 0.0012$. The Bayesian risk is $J_{SP}(k_{max}) = 0.132$ ($k_{max} = 1.01$) which is less than $I_{NH} = 0.141$ in the probability model.

4.4. *Human-machine Monitoring System*

Consider a human-machine system where the object is inspected when the system output is unknown state, X. Similar to probabilistic model, the expected loss of NHMS with inspection is expressed by using inspection cost C_I and probability of human error R, as follows;

$$J_{insp}(k) = \pi C_b Bel_D^k(S) + \overline{\pi} C_a Bel_S^k(D) + \pi(C_I + RC_b)m_k(X|_D) + \overline{\pi} C_I m_k(X|_S). \quad (17)$$

The optimal combination rule for the system is obtained by Prop 4.4.

Proposition 4.4
$$\min_k J_{insp}(k) = \begin{cases} J_{insp}(1) & \text{if } \frac{K+L}{C_I+KR} \geq 1 + \frac{q_S(S)}{q_S(D)} \\ J_{insp}(k_{max}) & \text{if } \frac{K+L}{C_I+KR} < 1 + \frac{q_S(S)}{q_S(D)} \end{cases}$$

Under fixed k, $J_{SP}(k) - J_{insp}(k) = (L - RK - C_I)m_k(X|_S)$ and $J_{FW}(k) - J_{insp}(k) = (K - RK - C_I)m_k(X|_D)$ so that D-S model leads to the Prop.4.5.

Proposition 4.5 Under fixed k, NHMS with inspection is more desirable than NHMS without inspection if and only if $min(K, L) - RK > C_I$.

Example 4 Consider again Example 2. $\frac{K+L}{C_I+KR} = \frac{20+0.999}{0.1+20*0.01} = 70.0$ and $1 + \frac{q_S(S)}{q_S(D)} = 1 + \frac{0.89*0.99}{0.01*0.11} = 802.0$ so that the Extra rule ($k = k_{max}$) is optimal by Prop. 4.4. Since $J_{SP}(k_{max})$ is the Bayesian risk as Example 3 shows, Prop.4.5 reveals that inspection is effective. Then, $J_{insp}(k_{max}) = 0.056 < J_{SP}(k_{max}) = 0.132$. Table 1 summarizes the results of our examples.

Table 1 Expected losses in Examples 1, 2, 3, and 4.

NHMS	Prob. model	D-S model
without Insp.	$I_{NH} = 0.141$	$J_{SP}(k_{max}) = 0.132$
with Insp.	$I_{insp} = 0.059$	$J_{insp}(k_{max}) = 0.056$

5. CONCLUSIONS

Non-homogeneous monitoring system that consists of an SP sensor and an FW sensor has two properties. One is to make the probability of conflict outputs low. The other is to identify the uncertainty zone in which sensors are hard to distinguish between safety and danger. This paper analyzes how to deal with the conflict outputs and the uncertainty zone via two models.

In the probabilistic model, it is shown that a machine system consider them as a safe state or as a danger state, respectively. A human-machine system can divide them into safety part and danger part by inspection activities. A condition is provided under which inspections by humans are effective in consideration with human error and inspection cost.

In the D-S model, it is established that the optimal combination rule partly divides two uncertainty domains into safety part and danger part and that the remaining part of the uncertainty zone is interpreted by the optimal system among SP system and FW system. Though they are performed by NHMS without inspection, our analysis has revealed that NHMS with inspection is more desirable than SP/FW system under certain conditions.

Acknowledgment This research was partially supported by Saneyoshi Scholarship Foundation.

6. REFERENCES

Henley, E.J. and H. Kumamoto (1985). *Designing for Reliability and Safety Control.* Prentice-Hall, New Jersey.

Inagaki, T. (1991). Interdependence between Safety-Control Policy and Multiple-sensor Schemes via Dempster-Shafer Theory, *IEEE Trans. Reliability*, Vol.R40-2, pp.182-188.

Klir G.J. and T.A. Folger (1988). *Fuzzy sets, Uncertainty, and Information*, Prentice-Hall, New Jersey.

Phillips M.J. (1980). K-out-of-n:G Systems Are Preferable, *IEEE Trans.Reliability*, Vol.R-29, No.2, pp.166-169.

Shafer G. (1976). *A Mathematical Theory of Evidence*, Princeton University Press.

Tanaka, K. (1994). Reliable Design of Safety Monitoring System Accounting for the Uncertainty Zone, Ibaraki Univ., Dept of Computer and Information Sciences, Technical Report 94-1.

Yager R. R. (1987). On the Dempster-Shafer Framework and New Combination Rules, *Information Sciences*, Vol.41, pp.93-137.

TOOD : TASK OBJECT ORIENTED DESCRIPTION FOR ERGONOMIC INTERFACES SPECIFICATION

A. Mahfoudhi, M. Abed, J-C. Angué

L.A.M.I.H., University of Valenciennes,
B.P. n° 311 59304 Valenciennes Cedex FRANCE
Tel. : + (33).27.14.14.61 E-mail : mahfoudhi@univ-valenciennes.fr

Abstract : Despite the recent progress in the domain of Man-Machine Interface engineering, several problems concerning the incompatibily between the information presentation to the user and his cognitive representation are still present. This paper presents a new Task Object Description methodology (TOOD). It tries to relate the characteristics of the user's task with those of the interface. The introduction of ergonomic concepts allows to take the human factors into account. And the joint use of the object oriented techniques and the High Level Petri Nets supplies complete, coherent and re-usable entities allowing to give a formal description of the interactive systems' characteristics and an appropriate specification of the user interface. An example, extracted from the air traffic control, is presented to illustrate this new methodology.

Keywords : Task Analysis, Man-Machine Interface, Specification, Co-operation, Object Oriented Techniques, Object Petri Nets, Human Factors.

1. INTRODUCTION

The technological evolution and increase of economic constraints have given birth to high-performance systems. However, several ergonomic problems are still present during the complex systems design where the presence of operators is and will remain necessary. These problems can be essentially related to the lack of taking into consideration the users' knowledge and human factors which are either implicit or explicit during the evaluation phase, and the absence of formal method for the Man-Machine Interface specification. So, it's very important to have models and methods that allow, on the one hand to make more accessible the users' knowledge, and mainly to make more formal and more detailed their description. On the other hand to allow the specification of the communication interface (Man-Machine Interface : MMI).

Over the last few years, the domain of man-machine interaction has been enriched with new methods and techniques. Some of them try to exploit the knowledge acquired by the cognitive theories in order to describe the operator tasks and its strategies (Richard *et al.*, 1992). The others use the simulation to study the decisional behaviour of the human operator (HO) (Cacciabue, 1988). At the same time, studies are carried out in order to modelize the man-machine interaction (Norman *et al.*, 1986) and to modelize the user interface (Coutaz, 1987) and (Green *et al., 1986*). Recently several users' tasks analysis and description methods, such as MAD proposed by Scapin and Pierret-Golbreith (1989) or the SADT/Petri method proposed by Abed (1990), have tried to describe the task as a set of operations and, at the same time, they integrate implicit information about human factors.

Our contribution consists in the proposition of a global methodology, called Task Object Oriented Description (TOOD), beginning from the task analysis and description to the MMI specification, based on the object oriented techniques and Object Petri Nets (OPN). TOOD has three goals : (1) to assure the transition between the different phases of the system development, (2) to allow the taking into consideration of the appropriate knowledge of the three system's components (the process or application, the MMI and the user), and (3) to offer a framework of efficient collaboration between the different contributors in the system development (ergonomists, computer specialists and users).

This article is structured into two sections : the first one presents a brief discussion of the tasks collection and modelling formalism used by TOOD, called "External Model". The second section explains the transition from the external model to the MMI specification called "Internal Model". The examples presented along this article are provided in the air traffic control context (Mahfoudhi and Abed, 94).

2. USERS' TASKS DESCRIPTION : "EXTERNAL MODEL"

Like the majority of task description methods, TOOD advocates the hierarchical approach for the tasks description. Although it's highly structured, the hierarchical approach allows to represent the user's cognitive model. Indeed, it allows to identify all the tasks to be carried out by the user with different choice of actions and possibilities of sequences.

Before presenting the different stages of the external model, it is advisable to define the different used terms. The generic term "Task-Object" indicates each task of the hierarchy. Indeed, the task-object is defined as an independent entity and responsible for a treatment, whatever his complexity level, to reply to a goal to be carried out with given conditions. Let us consider the air traffic control, "to plan the traffic" can be regarded as a task-object. In order to reduce its abstraction, this task-object can be decomposed into two children task-objects : "to plan the traffic in the sector" and "to plan the traffic in the sector's frontiers".

The task-object has a graphic form, inspired from the HOOD formalism (Michel, 1991) and Extended SADT method (Feller and Rucker, 1990), presented in Fig. 1.

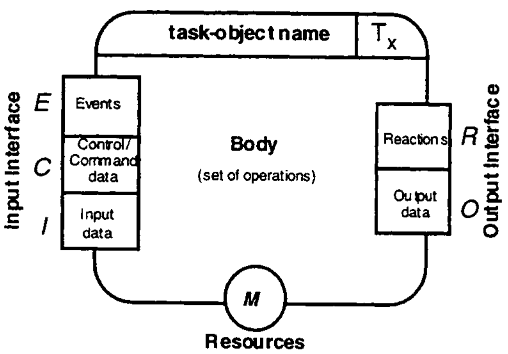


Fig. 1. graphic structure of the task-object

A task-object is also defined by a set of attributes, called "descriptors", that defines the execution conditions and the effects of the task, as well as the actions or sub-tasks to be carried out to reply to a given functional context.

Table 1. the task-object's descriptors.

Descriptor	Definition
Name	a name which identifies the task.
Input Interface	identifies the initial state of task-object. It is composed of three elements (E, C, I).
E	define the necessary Events for the task release.
C	Control/Command data : define the constraints to be respected at the time of the task execution.
I	Input data : define the list of data and information transformed by the task execution.
Output Interface	identifies the final state of task-object. It is composed of two elements (R, O).
R	Reactions : define the reports of the task-object execution (action realized or service demanded from other task-object).
O	Output data : defines the list of modified input data and/or a new data created after task-object execution.
Body	describes the operational model of the task-object.
Resource	defines the necessary human and material entities for the task-object execution.

The majority of methods presents the disadvantage that their descriptors are not exploited and their treatments are not detailed. TOOD has found a solution for this problem. Indeed, the notion of TCS (Task Control Structure) exploiting the task object's descriptors, allows to describe precisely the different task-object's states.

To establish the external model, three stages must be realized :

- task-objects identification,
- task-objects specification,
- description of the TCS (Task Control Structure).

2. 1. Task-objects identification

In order to identify all the tasks to be designed in the future system, TOOD begins by analyzing the existing system and the operator's current tasks. It allows to avoid the disadvantages of the existing system and to add the new desirable characteristics.

By a hierarchical decomposition, TOOD organizes the identified tasks-objects in a hierarchical tree form. It starts from the global task-object (the hierarchical tree's root) passing through the least abstract task-objets (the knots) and finishes with the terminal task-objects (the leaves).

Once all future system's tasks are identified, TOOD defines the constraints and relations between them. It also makes the distinction between the users' tasks and the computer's tasks. Indeed, it attributes all the interactive and manual tasks to the user, while the automatic tasks are attributed to the computer. Once the allocation bas been effected, TOOD takes an interest in the users' tasks because only this category can be used for specifying the users' interfaces.

2. 2. *Task-objects specification*

This specification stage defines all the execution conditions and the effects of each task-object. Based on the encapsulation concept, it describes what the task-object can do through the input/output interfaces.

For each task-object, the specification consists in :
- listing and identifying all the enclenchement events and the reactions. Indeed, the task execution can be asked for by the emission of an event, while the treatment report of the task-object is provided by a reaction.
- listing and identifying all the input, output and control/command data required and supplied by the task-object.

2. 3. *Task Control Structure : TCS*

The last stage describes the dynamic task-object behaviour. To that aim, a Task Control Structure "TCS" is used. The TCS is modelized by a Coloured Petri Net (Jensen, 1987) called "Petri Net Task Control Structure: PNTCS" (Fig. 2).

In order to remove any lack of determinism, two functions of data distribution (f and g) and a priority function (δ) are associated to the input and output transitions. Indeed, the functions f and g group together all the Pre and Post functions usually combined with the arcs of a coloured Petri net. Their aim is, on the one hand to select the necessary input and control/command data to activate the task-object with a given enclenchement event; on the other hand to specify the output data produced with the reaction. The priority function (δ) arranges the enclenchement events of the task according to their importance (alarms, interruption, temporal constraints, etc.).

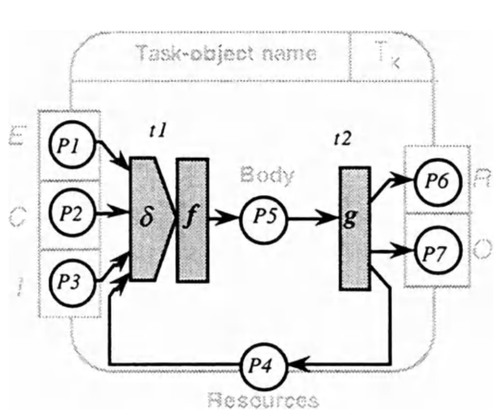


Fig. 2. Task Control Structure (TCS)

The formal aspect of the TCS allows to identify, at any time, the current state of the task-object (treatment authorised, waiting for an enclenchement event, producing a reaction, etc....).

The resulting document of the external model includes two kinds of description : a graphic specification for a clean, legible and exploitable representation, and a textual specification for a complete description (Mahfoudhi *et al.*, 94).

3. USER INTERFACE SPECIFICATION : "INTERNAL MODEL"

The aim of this stage is the automatic passage from the users' tasks description (external model) to the MMI specification (internal model) Fig. 3. It completes the external model by defining the operational level of each terminal task-object. Indeed, it allows to define all the necessary action plans and manipulated objects for the task executing. So, the resources of each terminal task-object become its component-objects belonging to the classes : Interface, Application and Operator (Fig. 3).

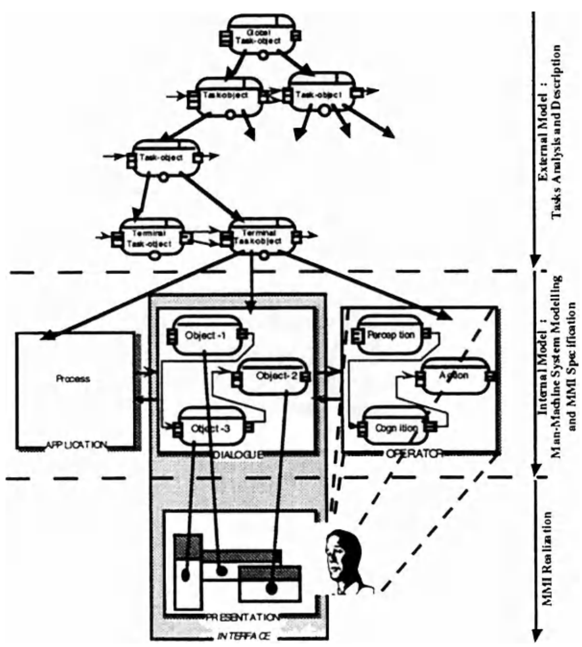


Fig. 3. From the users' tasks description to the MMI specification.

The MMI specification is carried out on two stages. The first one consists in specifying of the component-objects of each terminal task-object, with an aggregation of these component-objects; the second stage allows to specify the user interface.

3. 1. Component-objects specification

All the component-objects co-operate in a precisely manner in order to fulfil the aim of the terminal task-object in reply to a given functional context. A component-object shall be defined from its class (Interface or Operator) and provided with a set of states and a set of operations (or actions) which allow to change these states. For example, from the P3 state (strip selected) of the component-object "new strips table" the operator has the possibility to carry out two actions : t3 (open a road-zoom) or t5 (temporize the new strip) Fig.4. On the other hand, the set of states and operations of an Operator component-object represents the different possible procedures for the execution of the terminal task. Indeed, the procedure represents the different activity phases of a human operator : situation apprehension, goals identification, preparation of an action plan, application of this action plan, control of the situation, correction (Norman *et al.*, 86).

Graphically, the component-object is presented in an identical structure that the one of a task-object in the external model. However its internal control structure called Object Control Structure "ObCS" is modelized by an Object Petri Net "OPN" (Sibertin, 85).The OPNs are characterized by the fact that the tokens which constitute the place markings are not atomic nor similar entities, but they can be distinguished from each other and take values allowing to describe the characteristics of the system.

In addition to its formal aspect, the ObCS enjoys a simple and easily understandable graphic representation, allowing to represent - with the places of the OPN - all the possible states of the component-object, and with the transitions, to represent all the operations and actions that can be taken from these states. The graphic representation used for the ObCS is inspired by the cooperative and interactive objects formalism proposed by Palanque (1992).

The different states of the operator component-object come down to the three states : Perception, Cognition and Action. The ObCS allows the distinction between these states (table 2).

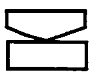
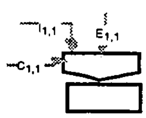
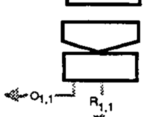
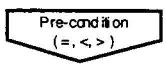
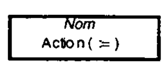
Table 2. the states of the operator component-object

Place	State
	perception
	cognition
	action (an operator action on the interface)

The communication between the component-objects is carried out through their input and output interfaces. So, an action "A" executed by a component-object "X" (operator) on the component-object "Y" (interface) can be read as : the component-object "X" executes his reaction operation corresponding to his quiry of the action A. This execution is rendered by a reaction R in the output interface of the component-object X. The output interface transmits the reaction R to the input interface of the component-object Y. So the reaction R becomes an event E. And lastly this event activates the service operation of the component-object Y corresponding to the action A asked by the component-object X.

Graphically, the ObCS allows the distinction between the different kinds of operations (private operation, service operation and reaction operation) table 3.

Table. 3. the different transitions in the ObCS

Transition	Description
	**tP**: private transition does not communicate with the interfaces of the component-object.
	**tS**: service transition. The Ei, Ci and Ii communicate with the input interface.
	**tR**: reaction transition. The Ri and Oi communicate with the output interface.
	The Pre-condition part of the transition which has a test about the input data of the transition.
	The action part of the transition which has the name and the expression of the operation

The execution of a service (service operation) by the component-object is equivalent to the fire of the service transition (tS) associated to this service.

The fire of a reaction transition (tR) renders the fact that the component-object executed the reaction operation (produced a reaction) associated with this transition. The ObCS can also contain transitions that are not associated to any services nor to a reaction; these transitions are called "Private Transition" (tP). They modelize the internal changes of the states of the component-object.

An example taken from the air traffic control, corresponding to the terminal task-object "take knowledge the new flight" taken from (Mahfoudhi and Abed, 94), needs using two component-objects : "a New Strips Table : NST" and " Organic Controller : OC" (figure 4). The comportment of the component-object "a New Strips Table" is defined by four states P1, P2, P3 and P4. From each state the Organic Controller can carry out a group of actions (transitions). From the P3 state (strip selected), For example, he has the possibility to achieve two actions : t3 (open a road-zoom) or t5 (temporize the new strip).

For the component-object "Organic Controller", the set of states and operations represents the different possible procedures to execute the terminal task "Take knowledge of a new flight" in reply to a given functional context. So, the display of a New Strip NS in the component-object "new strips table" invokes, by the event E2,1, the operation service "Consult the NS" of the component-object "Organic Controller OC". According to his selection "Ch=", the organic controller carries out a first reading of the NS information ("Consult the road" or "Consult the level"). After this reading, he changes his state into cognition in order to evaluate his information level. Then he decides to "read again the basic information" or "to ask for additional information". The asking for additional information expresses itself by a change of his state into "Action" in order to "select the NS" and to "open the Road-Zoom". Both actions transmit R2,2 and R2,3 reactions to the component-object "new strips table". It is to be noticed that the organic controller carries out the action "open a road-zoom" only after receiving the event E2,2 confirming that the action "Select the NS" has been carried out. Once the Road-Zoom has been opened, the Organic Controller changes his state into "information reading" in order to read the additional information and then into the "situation evaluation" state to decide either to read again the information, or "to temporize the NS" or to invoke the terminal task-object "T112 : analyse the entrance conditions".

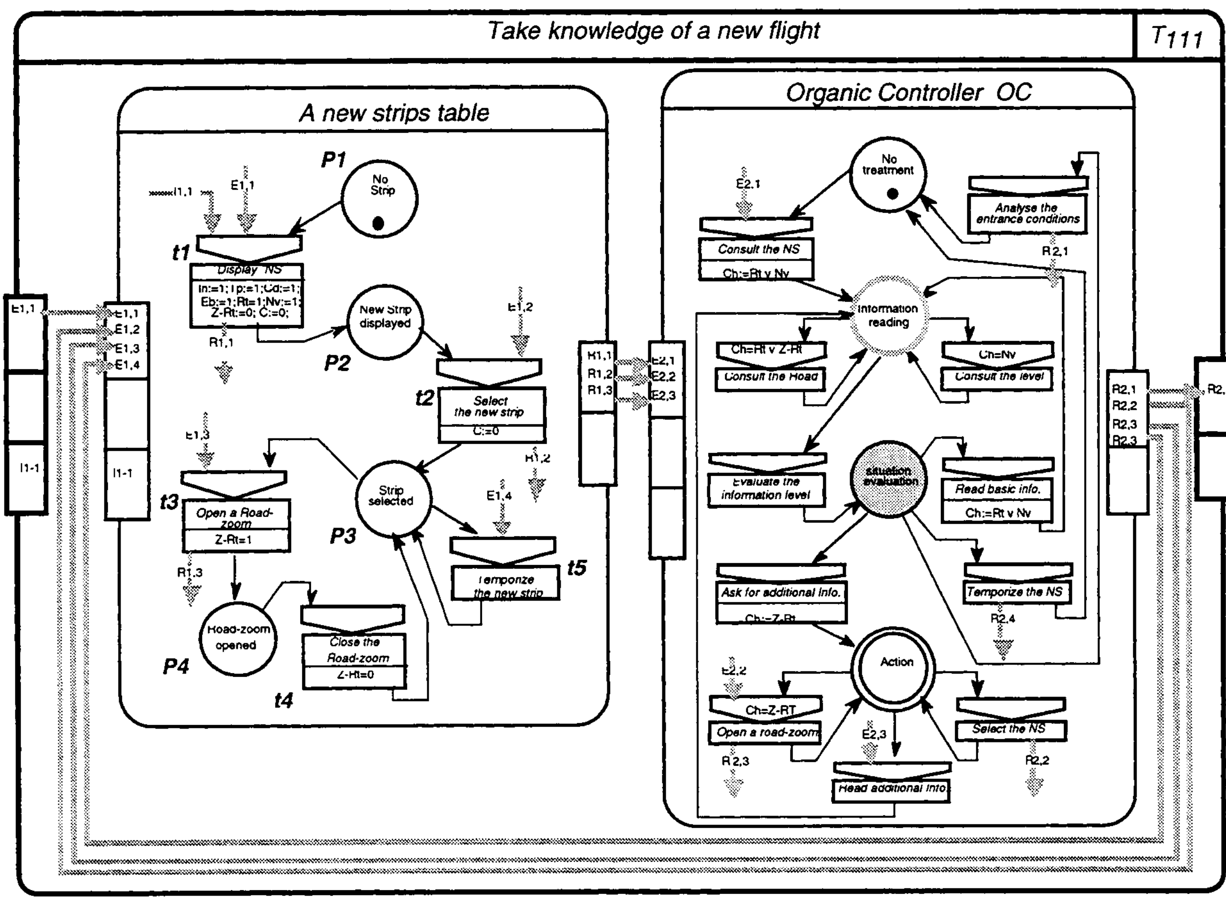


Fig. 4. A graphic Specification of the component-objects "New Strips Table" and "Organic Controller"

3. 2. Aggregation mechanism

In order to realize the MMI in its real structure, the construction of the object classes of the MMI suggests the aggregation of the different component-objects which have the same name, specified during the description of the internal model of each terminal task-object. This aggregation mechanism is comparable to the composition relation of the HOOD method called the parent/child relation.

Thus, an object class of the MMI is built according to the duplication of all the elements (events, control/command data, input data, reactions, output data and ObCSs) of the component-objects which have the same name. Yet, there would be some modifications which are detailed in (Mahfoudhi and Abed, 94).

4. CONCLUSION

The TOOD methodology enjoys the contributions of methods and concepts taken from cognitive sciences and ergonomy domains together with those of the software engineering domain. It provides a framework of efficient collaboration between various users and between ergonomists and computer specialists. Its formalism allows on the one hand to define in a formal, coherent and structured way, the different entities intervening in the task model, and on the other hand to specify an adapted interface to the users' characteristics. Moreover, its mathematical formalism allows it to have a tool for the validation and the simulation. There is still to develop a language leading to its exploitation on a large scale.

5. REFERENCES

Abed, M., 1990, Contribution à la modélisation de la tâche par des outils de spécification exploitant les mouvements oculaires: application à la conception et l'évaluation des Interfaces Homme-Machine [contribution to the task modelling with a specification tools exploiting the ocular activities : application to the man-machines interfaces design and evaluation], Doctorate thesis, University of Valenciennes France.

Cacciabue, 1988, Modelling Human Behaviour in the context of a simulation of Man-Machine Systems, Trainaing Human Decision Marking and Control, Elsevier Science Publishers B.V., North Holland,

Coutaz, J., 1987, PAC : An Implementation Model For Dialog Design. Interact'87, pp 431-436, Stuttgart.

Feller, A., and Rucker, R., 1990, Extending Structured Analysis Modelling with A.I.: An Application to MRPII Profiles and SFC Data Communications Requirements Specifications, IFIPS Conference.

Green, 1986, A survey of three dialogue models. ACM Transaction on graphics. **Vol. 5, Nb. 3** pp. 244-275.

Jensen, K., 1987, Coloured Petri Nets, In Petri Nets : Central models and their properties, LNCS **Nb. 254**, Spring Verlag.

Mahfoudhi, A. and Abed, M., 1994, Description orientée objet des tâches du contrôleur pour la spécification et la conception des interfaces [a controler's tasks object oriented description for the interfaces specification and design], Contract research report CENA 94, University of Valenciennes France.

Michel, L., 1991, Conception orientée objet : Pratique de la méthode HOOD [object oriented design : practice of the HOOD method], Dunod Press Paris.

Norman, D. and Draper, 1986, User centered system design, Lawrence Erlbaum Associates, Publishers.

Palanque, P, 1992, Modélisation par objets Coopératifs Interactifs d'interfaces homme machine dirigées par l'utilisateur [modelling of the man machine interfaces managed by the users using interactive and cooperative objects], Doctorate thesis, University of Toulouse I France.

Richard J.F., Poitrenaud S., Tijus C.A., Barcenilla J., 1992 Semantic of action networks : a cognitive model for human and system interaction. HCI 92, Université de Paris 8.

Scapin, D.L. and Pierret-Golbreith, C., 1989, MAD : Une Méthode Analytique de Description de Tâches [MAD : a tasks analytic description method], Actes du Colloque sur l'ingénierie des Interfaces Homme-Machine, Sophia-Antipolis, France.

Sibertin, B.C, 1985, High-level Petri nets with Data Structure. 6th European Workshop on Petri Net and application, Espoo, Finland.

EVALUATION OF TWO HUMAN OPERATOR MODELS OF THE NAVIGATOR'S BEHAVIOUR

Robert Papenhuijzen and **Trijnie Dijkhuis**

Ministry of Transport, Public Works and Water Management, Transport Research Centre (AVV), P.O. Box 1031, 3000 BA Rotterdam, The Netherlands

Abstract: A computer program to simulate the behaviour of the complete system of a ship under human control will make a powerful tool for investigations in the field of navigation and fairway design. Hence, two models have been developed to simulate the behaviour of the navigator, the human in charge of the navigation process. On the basis of a large number of ship bridge simulator experiments, the two models have been evaluated, and compared one with the other.

Keywords: fuzzy control, human factors, human supervisory control, linear optimal control, ship control, simulation.

1. INTRODUCTION

Fairway dimensioning and safety problems are, in fact, as old as navigation itself. Finding adequate solutions to such problems is gaining more and more importance. Larger vessels are built, being more difficult to handle, vessel traffic intensifies, and there is an ever growing number of ships that transport dangerous cargo. All this, in addition to the increasing need for safe, cost effective and environmentally acceptable alternatives for road transport, has given new impulses to the research in the field of fairway design and regulation.

Since the early seventies, this type of research is generally performed using a ship bridge simulator, for accomplishing realistic sailed tracks, given a combination of a professionally steered vessel and any — generally confined — hypothetical or existing waterway. But, in fact, many highly relevant studies are carried out only partially, or are not carried out at all, due to the high cost of operating a ship bridge simulator. Therefore, it was decided to venture the development of a simulation program that covers all elements of the navigation process, including the human operator component, referred to as the *navigator*. A suitable model of the human navigator not being available, the development of such a model was taken to hand within the framework of the *navigator project*.

As a result, a model of the state estimation behaviour of the navigator has been proposed, based on Kalman filtering. A simple prototype was built in order to demonstrate its feasibility (Papenhuijzen, 1988). Next, due attention was paid to modelling the planning and control behaviour. For this, two basically different approaches were chosen, one rooted in linear optimal control theory (Papenhuijzen and Stassen, 1987, 1989), whereas the other one was based on fuzzy set theory (Papenhuijzen and Stassen, 1992). Both approaches have yielded an operational navigator model, that was coupled with suitable models of an environment and a ship.

In this paper, the two models are discussed only briefly. Emphasis is put on the validation of the models, based on the comparison between model output and a large number of ship bridge simulator experiments. For this, experiments have been performed with two different inland vessels and two different sea-going vessels, under various environmental conditions. Next, the models are evaluated comprehensively, considering current fields of application, as well as the need and the possibilities for future extensions.

2. OVERVIEW OF THE MODELS

As indicated before, two basically different navigator models have emerged from the project, differing in the way in which the navigator's planning and control behaviour is modelled. For the first model, planning and control simulation is based on linear optimal control theory, whereas the other one employs in total ten Fuzzy Logic Controllers (Lee, 1990) in order to simulate the various decision making processes. Both models may either include a Kalman filter to simulate state estimation, or draw on perfect state information upon which to base planning and control actions.

The working principles of the planning and control submodels are summarized in the remainder of this section. An in-depth treatment of the two different approaches is given by Papenhuijzen (1994).

2.1 The control theoretic approach.

Conceptually, the complete planning process may be considered to comprise both so-called long term planning and short term planning. Long term planning is aimed at determining a desired route and a preliminary time planning, whereas short term planning decides on the exact trajectory to follow. Long term planning is, in fact, the voyage preparation process, yielding just a few target states, corresponding time instants, and an indication of how important it is to realize the various components of the target states. Given the field of application of the program, it was not considered necessary to simulate long term planning as such. As an alternative, a very small number of target states, generally one, is to be specified by the user, in order to define the navigation task.

The problems of track planning and track following are basically interrelated. In defining the short term track to be followed, not only the attractiveness of a track in terms of acceptable risk, but also the dynamical properties of the ship have to be considered. For that reason, it has been found logical to develop a submodel in which the description of the navigator's short term planning behaviour and track following behaviour are integrated.

The ultimate goal that has to be achieved by the track planning and track following submodels, includes the following aspects. In the first place, a target state, an outcome of the long term planning process, has to be reached within a given span of time. Furthermore, the risk that is involved in sailing to the corresponding position, and arriving there with a given velocity, is to be kept sufficiently small, as are the necessary control inputs. Linear optimal control theory provides an elegant framework for generating realistic control inputs on the basis of the combined subgoals mentioned. The three distinctive aspects of short term planning reappear in a control theoretic context as terminal error weighting, state weighting and control effort weighting, respectively.

Unfortunately, deriving an optimal control strategy, in terms of a linear optimal control law, is not straightforward. In this particular case the optimal solution of the regulator problem, which includes the short term track to be sailed, is not known. On the contrary, the optimal combination of a trajectory and a control signal should be a *result* of the algorithm, rather than an *input* for it. Consequently, the control problem to be solved is the so-called generalized discrete-time linear optimal regulator problem (Papenhuijzen, 1987).

Solution of this control problem is achieved by iterative application of the linear optimal control law, which is derived, respectively, around an anything but optimal initial solution to start with, and around the latest improved solution that has resulted from applying the control law. As soon as the algorithm has duly converged, the latest control law is saved to be used to simulate track following, featuring multi-variable system control, ie vessel control by applying both rudder and engine speed variations.

2.2 The fuzzy set approach.

Most navigators use to interpret and plan their course on confined waterways as a concatenation of lanes and circle shaped bends. Hence, the planning model is designed to supply the track following model with a desired track in terms of straight lines and arcs of circle. Apart from that, the planning model yields a — fuzzy — definition of the navigator's perception of safe manoeuvring zones, in order to enable the track following model to assess the safety consequences of deviations from the ideal track. Additionally, the safety zones definition is also used by the planning model itself, since track planning is impossible without an appropriate representation of safety.

Supervisory and manual control behaviour, the execution of the travel plan, is simulated by the track following submodel. Every second, future states are predicted. The predicted states are evaluated by relating them to the perception of safety as determined by the track planning submodel, and by assessing the measure to which future states diverge from the desired states as defined by the planned track. If necessary, a control action is carried out.

Contrary to what is the case for the control theoretic model, the control actions taken by the fuzzy set navigator model comprise only rudder commands, and not engine settings. Constant engine speed setting is assumed during a whole run.

3. VALIDATION EXPERIMENTS

Validation and evaluation of the models has been accomplished by comparing the outcomes of a large number of ship bridge simulator experiments to the corresponding outcomes of the two navigator models. It was considered essential to be able to judge the performance of the navigator models in inland navigation situations as well as in confined sea navigation situations. Another important issue to be studied, was the suitability of the models for dealing with different types of ships. Further, the ability of the models to cope realistically with environmental influences, wind and current in particular, had to be investigated.

Consequently, two fairways, both located in the Port of Rotterdam area, were selected. For the sea navigation experiments, the approach to Mississippi harbour was simulated. The inland navigation experiments were performed on the Hartel canal, the route from the village of Geervliet westward to the Rozenburg lock. For the first fairway, a moderate wind and a heavy wind condition was selected. For the inland navigation experiments, apart from a simple reference condition, a heavy wind condition was defined, as well as one with relatively strong current from behind.

For both the sea navigation trials and the inland navigation trials, a relatively small ship type was used, and a relatively bulky one. As a small sea vessel, a loaded 210 m long container carrier was employed. The heavier deep sea vessel was a loaded 290 m long bulk carrier. For the inland experiments a relatively small loaded tanker was used, and, as the more difficult vessel to handle, an empty six barge push tow unit.

Fortunately, it was possible to employ all highly experienced subjects for the experiments. Four harbour pilots cooperated on the sea navigation trials, each one of them performing four runs per condition. Two inland tanker captains carried out the inland navigation experiments with the tanker, both conducting a share of five runs per condition. To the experiments with the push tow convoys, two push tow captains contributed, each one also performing five runs per condition. As a result, a total number of 124 trials were carried out, divided into 10 experiments. Additionally, the same experiment program was carried out with both the control theoretic simulation model and its fuzzy set counterpart.

4. AN EXAMPLE

In this section, as an example, some results of one of the sea navigation experiments, are presented. For this purpose, the experiment with the bulk carrier under moderate wind conditions was selected.

Fig. 1.a shows the results of one of the 16 trials of the experiment, performed by a pilot on the ship bridge simulator. To the left, the track of the ship is plotted, as well as the hull of the ship, two times per minute. To the right, a time plot of the engine setting and the rudder setting is given. This example demonstrates the way in which mariners use to execute a manoeuvre like this. Major course changes are initiated with a generous rudder shift, which is maintained until the ship has noticeably started to respond. Keeping the rate of turn constant at a certain level then, is achieved by setting a somewhat smaller rudder angle, until the turn is counteracted by applying a rudder angle in the opposite direction. Course keeping on more or less straight tracks is realized by relatively frequent application of relatively small rudder adjustments. Further, a minor engine speed variation is executed when starting the first turn. A more significant engine speed adjustment is applied to help counteracting the second turn.

Results of a trial of the same experiment, carried out with the control theoretic navigator model, are shown in Fig. 1.b. Obviously, an optimal control strategy would result in perfectly smooth tracks and control signals, if it was not for a shaping filter to modify the calculated optimal control settings. However, large deviations between the optimal trajectory and the realization are not allowed, as a consequence of which the control signals in Fig. 1.b. are clearly a discrete version of an otherwise almost identical pair of smooth signals. Unfortunately, it may be clear that the character of the control signals generated by the pilots is anything but represented by the optimal control navigator model. On the other hand, the ability of the model to cope with wind and current disturbances is quite remarkable, and enables the model to stay well within the limits that are set to the validity of a linear control approach. Furthermore, in spite of the unusual rudder control signal, the resulting track seems realistic enough. Another interesting point is the way in which the model manages to

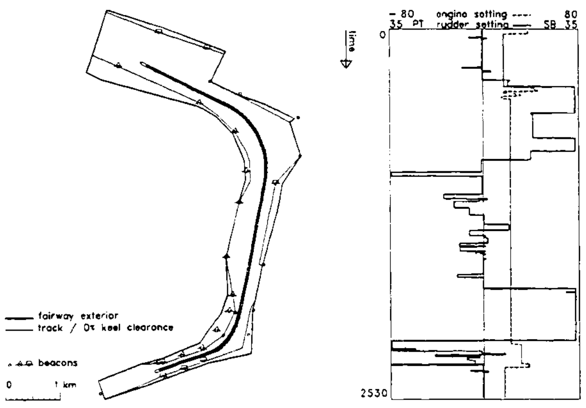


Fig. 1.a. Simulation of an approach with a bulk carrier to Mississippi harbour (southward), as performed by a harbour pilot on the ship bridge simulator.

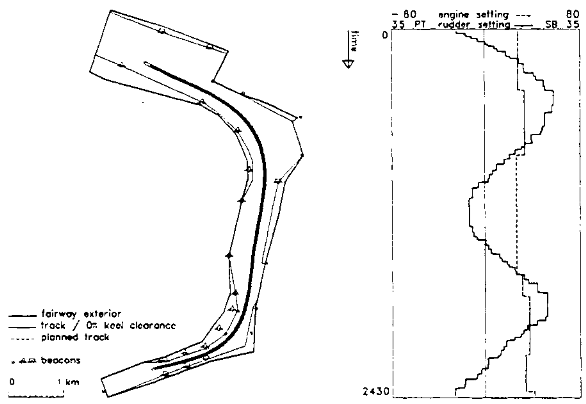


Fig. 1.b. Simulation of an approach with a bulk carrier to Mississippi harbour (southward), as performed with the control theoretic navigator model.

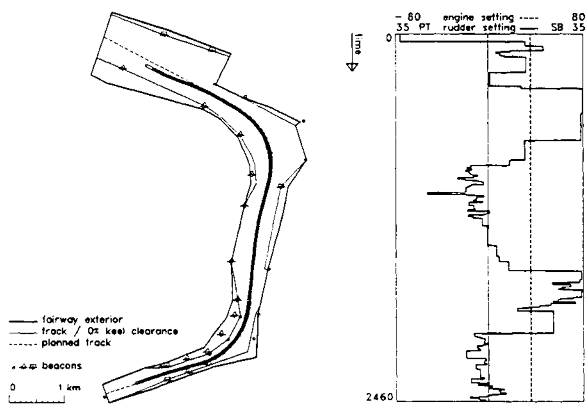


Fig. 1.c. Simulation of an approach with a bulk carrier to Mississippi harbour (southward), as performed with the fuzzy set navigator model.

combine engine speed variations and rudder control, navigating the ship apparently effortless to its destination.

A similar trial, performed by the fuzzy set navigator model, is presented in Fig. 1.c. It may be clear that here the character of the rudder signal comes much closer to reality than is the case when applying the control theoretic model. This concerns sailing a bend as well as lane keeping. As indicated before, the model does not feature engine speed control, which, apparently, does not prevent it from coping adequately with environmental disturbances, and realizing an acceptable track.

In order to be able to judge, for all trials that belong to an experiment, the model performance with respect to track variation and control effort aspects, also a second output format was used, as indicated in Figs 2.a to 2.c. To the left in the figures, the total fairway space that was occupied by any ship, the so-called swept path, is indicated by dashed lines. An impression of the distribution of all tracks that relate to a given experiment is given by drawing the individual tracks.

Further, to the right, the results of analyzing the corresponding rudder setting signals are shown. The curved solid line represents as a function of the distance accomplished by the ship, running top to down: the so-called average mean value of the rudder settings, $\mu_{av}(s)$, the definition of which is clarified in Fig. 3. In this figure, also the definition of the distance dependent average standard deviation of the rudder setting $\sigma_{av}(s)$ is given. The dashed lines in the right part of Figs 2.a to 2.c indicate the band of $\sigma_{av}(s)$ around $\mu_{av}(s)$. In this way, by $\mu_{av}(s)$ an idea is given of the tactical decisions that underlay the manoeuvre, whereas $\sigma_{av}(s)$ serves as an indicator of the control effort spent, all along the manoeuvre. Thus, as argued by Schuffel (1986), the complexity of the various parts of the manoeuvre can be judged, as well as the potential risk that is incurred.

The relationship between the left and the right part of the analysis output is indicated by drawing cross lines at a nominal spacing of 200 metres, the lines in the track plot corresponding one to one with those in the control signal diagram.

Figs 2.a to 2.c show the analysis plots that correspond to the experiments to which the trials in Figs 1.a to 1.c belong, based on all 16 runs making up each experiment. Here, a value of 50 m was used for the nominal distance between the cross sections in which the analysis steps were performed, and the time window T was 90 seconds.

As to the control theoretic model output, from Figs. 2.a and 2.b again the tentative conclusion can be drawn that fairly realistic tracks are produced. On the other hand, the corresponding rudder signal analysis plots appear to be perfectly ruthless when

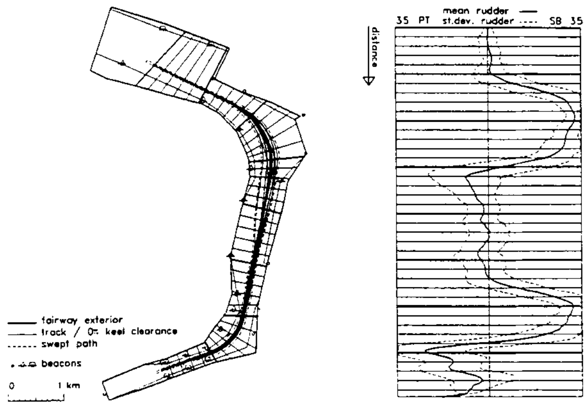


Fig. 2.a. Analysis of all trials of an experiment concerning the approach with a bulk carrier to Mississippi harbour (southward), as performed by harbour pilots on the ship bridge simulator.

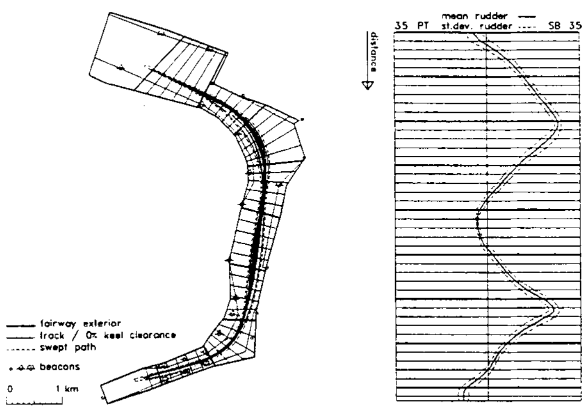


Fig. 2.b. Analysis of all trials of an experiment concerning the approach with a bulk carrier to Mississippi harbour (southward), as performed with the control theoretic navigator model.

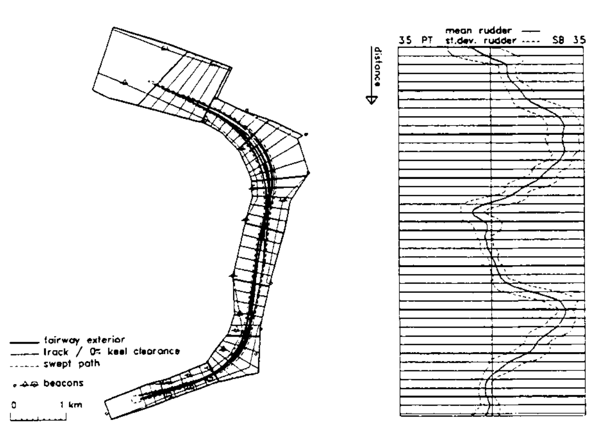


Fig. 2.c. Analysis of all trials of an experiment concerning the approach with a bulk carrier to Mississippi harbour (southward), as performed with the fuzzy set navigator model.

it comes to distinguishing between the performance of this mathematical concept and the behaviour of a real navigator.

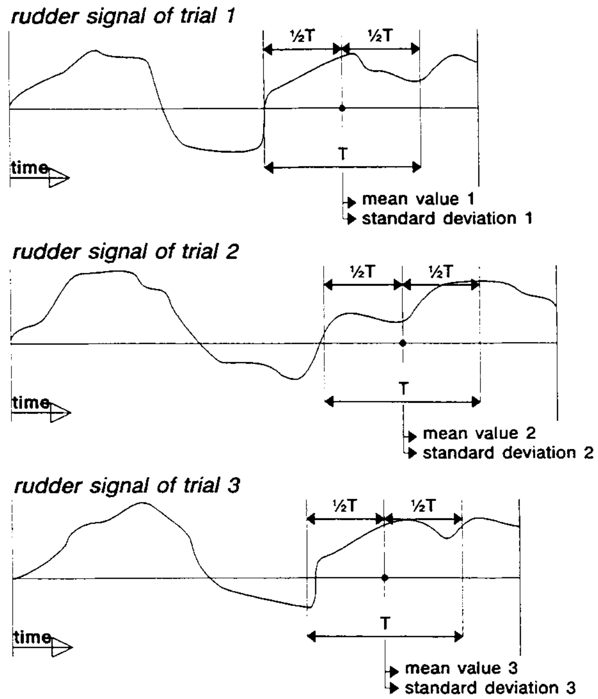


Fig. 3. Principle of the analysis of the rudder control signals. Given an elapsed distance, for every trial the corresponding time instant is determined, indicated on the time axes by the bold dots. Then, for all signals, three in this example, the mean value and the standard deviation of the rudder setting within the time window T is calculated. The average of the mean values 1 to 3 is the average mean value $\mu_{av}(s)$ which is plotted in the analysis plots, as a function of the position along the axis of the fairway. Identically, the average standard deviation $\sigma_{av}(s)$ is calculated as the average of the three indicated standard deviations.

By comparing Fig. 2.c to Fig. 2.a, it can be concluded for the fuzzy set model that the tracks itself, as well as the apparent variation in track shapes, closely approximate reality. The conclusion that was drawn in the foregoing, namely that on the whole the navigator's behaviour as to rudder control is much better represented by the fuzzy set model than by the control theoretic model, is firmly ascertained here.

5. DISCUSSION OF THE RESULTS

From the experimental results, it was concluded that both approaches, the control theoretic one and the fuzzy set one, have resulted in a practical model that yields realistic ship tracks, considering only single ship situations. There are some important differences between the two models, however, concerning either the functionality of the models as they are existent now, or the possibilities for future enhancement of their applicability. In the following, this is examined in more detail, discriminating between seven major issues.

performance — Neither of the models appears to pose serious problems with respect to the necessary computing time. Generally, both models run about twice as fast as real time on a 50 Mhz 486 PC.

basic philosophy — The basic philosophy underlying the control theoretic approach, is that explicit modelling of the navigator's behaviour is achieved only on the level of the navigator's control objectives. The resulting control actions are derived from that, applying some plausible mechanism, which does not necessarily ensure perfect resemblance of the model output to the individual actions of a real navigator. The performance of the system to be controlled, however, may just as well be quite realistic. Contrarily, the fuzzy set model is a descriptive human operator model, featuring more explicit representation of individual decisions made by the navigator. Consequently, the output of the fuzzy set navigator model is much more realistic, which is essential when for instance safety studies are carried out, focusing on control effort assessment.

track planning strategy — The two models differ substantially in the way in which track planning is accomplished. The control theoretic model determines the optimal track on the basis of weighted minimization of risk, control effort and terminal error. On the one hand, the result is a highly generic method, which, in principle, enables the model to cope even with such complex navigation tasks as docking. On the other hand, this technique excludes the possibility for tracks that consist only of lanes and arcs of circle to emerge from the planning process. From the experiments it can be concluded, however, that generally the planned tracks of the control theoretic model look quite realistic. Only if unconditional representation of the real planning process as performed by, for instance, a harbour pilot is imposed, this method may be considered to lack realism. In those, probably rare instances, the fuzzy set model will be superior. However, the cost would have to be paid for this, is some loss of universality, as for certain complex manoeuvres it could be too much of a simplification to consider the planned track as a concatenation of a limited number of lanes and arcs of circle.

engine speed control — Originally, it was decided that in this stage it is not yet imperative to include engine speed setting in the navigator models. Eventually, however, this should be implemented after all. For the control theoretic model this has been achieved already, since engine speed control is inevitably inherent in the control theoretic implementation. Contrarily, extension of the fuzzy set model into a navigator model that includes engine speed control as well as rudder control is much more complicated. The point is that it is not sufficient to modify only the track following submodel. Engine speed setting at a given moment should not only be concerned with the particular manoeuvre that is performed then. Much more important is the general plan about where to be at what instant, and what speed to possess then, thus anticipating the manoeuvres to come. As a consequence, the fuzzy set model can only be made to feature engine speed control at the cost of considerable modification of the entire model.

time-varying environment definition — If the experiment involves navigating tidal waters, and if the trial covers significantly more than, say, a quarter of an hour, a time dependent definition of the fairway geometry and current and wind influences has to be employed. This is something that requires no basic extensions to the control theoretic model, since the linear optimal control law is principally time-varying, as are the corresponding weighting parameters. Contrarily, in the planning process of the fuzzy set model there is no direct relationship between time and the resulting planned position path. An attempt could be made to compensate for that, but the success of such a modification is doubtful.

traffic simulation — Eventually, it should be possible to cross the single ship barrier, and to develop a simulation system that includes more than one ship, featuring realistic interaction between the individual navigators. In its simplest form, only one ship would be equipped with a navigator model, the other ships following predefined tracks. This so-called floating island concept is fairly simple to implement for the control theoretic approach. It means that the risk definition has to be based on time-varying effective fairway boundaries, which, as discussed in the above, is not a serious complication. If, however, the other ships are supposed to react to the own ship or one to another, the situation becomes more complicated. In that case, the state space has to be extended so as to include the state variables of all individual ships, in order to realize an overall optimum. For the fuzzy set model it will be more difficult to be extended so as to include floating island traffic simulation. As stated before, here, the planned track lacks a direct relationship with time, which is essential for considering a moving object as part of the physical environment. In fact, full interaction would be even easier to realize, as this is an extensively explored field already, given the work that has been devoted worldwide to collision avoidance systems.

other applications — Apart from looking at the suitability of the various concepts for modelling the behaviour of the navigator, it is interesting to consider whether there are entirely new uses for the navigator models that have been developed. The most obvious, and at the same time most daring one, is the use of a navigator model for automatic guidance of a real ship. Neglecting for the time being such important issues as legislative restrictions and economical relevance, it is felt that automatic navigation under some conditions need not be far off, given the results of this project. In principle, automatic navigation in single ship situations could be performed directly by coupling a navigator model to standard positioning and course measuring equipment, and supplying adequate chart, wind and current data. If encounters with other vessels have to be dealt with, a RADAR-based recognition and prediction system must be available. In addition, major extensions to the navigator model have to be accomplished then, analogously to what was discussed in the previous paragraph. The suitability of either one of the navigator models for automatic navigation purposes can be determined on the basis of the following considerations. To start with, the conclusions of the discussion on traffic simulation are particularly relevant here. Secondly, in the case of automatic navigation it is not enough to employ methods that cover most, but not all possible situations. As a consequence, particularly based on the second consideration, it is concluded that for automatic navigation developments the control theoretic model is probably the best option. This conclusion is supported by the fact that in the experiments the control theoretic model managed to steer the vessels extremely accurately along the planned tracks, even though:

- there were, deliberately, considerable mismatches between the hydrodynamic model of the real vessel and the one residing in the navigator model as an internal model of the ship's behaviour;
- considerable environmental disturbances were acting on the real ship.

In order to illustrate this, Fig. 4 shows the result of a trial with the push tow unit, steered by the optimal control model, subject to a heavy south-west wind.

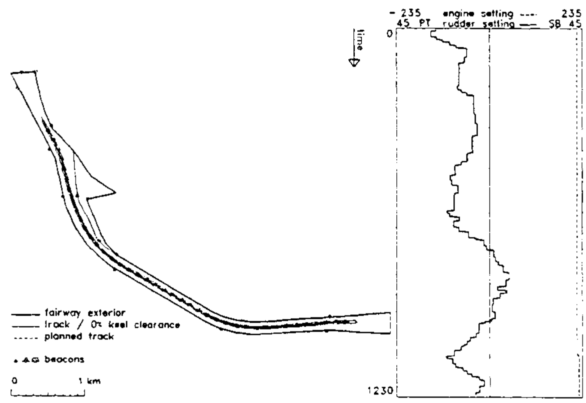


Fig. 4. Simulation of a manoeuvre with an empty six barge push tow unit on the Hartel canal (westward), with heavy south-west wind, as performed with the control theoretic navigator model.

6. CONCLUSION

On the basis of a large number of experiments, it was concluded that the fuzzy set model is the better one when it comes to realistically simulating the way in which a real navigator accomplishes rudder control. The control theoretic approach is probably superior as to developing into a system that includes complex planning behaviour, with significant engine speed variations and complex collision avoidance manoeuvres. If, within the framework of such a project, realistic rudder control is essential, the idea of combining the two concepts suggests itself. Apart from that, it is recommended to probe the development of a sophisticated navigation system, based on the control theoretic navigator model, that features not only automatic track following, but also automatic track definition.

REFERENCES

Lee, C.C. (1990). Fuzzy Logic in Control Systems: Fuzzy Logic Controller — Part I & II. *IEEE Trans. Systems, Man and Cybernetics,* **SMC-20,** 404—435.

Papenhuijzen, R. and H.G. Stassen (1987). On the modelling of the Behaviour of a Navigator. *Proc. 8th Ship Control Systems Symposium,* The Hague, 2.238—2.254.

Papenhuijzen, R. (1988). On the Modelling of the Behaviour of a Navigator; Background and Proposal for a Control Theoretic Approach. In J. Patrick and K.D. Duncan (Eds.), *Training, Human Decision Making and Control.* Elsevier (North-Holland), Amsterdam, 189—203.

Papenhuijzen, R. and H.G. Stassen (1989). On the Modelling of Planning and Supervisory Behaviour of the Navigator. *Proc. 4th IFAC/IFIP/-IFORS/IEA Conference on Man-Machine Systems,* Xi'an, China, 19—24.

Papenhuijzen, R. and H.G. Stassen (1992). Fuzzy Set Theory for Modelling the Navigator's Behaviour. *Proc. 5th IFAC/IFIP/IFORS/IEA Conference on Man-Machine Systems,* The Hague, 2.1.4.1—2.1.4.7.

Papenhuijzen, R. (1994). Towards a Human Operator Model of the Navigator. Ph.D. Thesis, Delft University of Technology, Delft.

Schuffel, H. (1986). Human Control of Ships in Tracking Tasks. Ph.D. Thesis, TNO-IZF, Soesterberg.

OBSERVATION IN MARITIME EMERGENCY MANAGEMENT

Torkil Clemmensen

Danish Maritime Institute
Hjortekærsvej 99
Dk-2800 Lyngby
Denmark
Tel: +45 45879325
Fax: +45 45879333
E-mail: clemmensen@axp.psl.ku.dk

Abstract: Observation is part of the work done by the senior officers onboard a merchant ship when a fire or collision has occurred. This paper describes the different observer positions and the technological support available to the captain in a simulation exercise. An analogy to scientific observation accentuate that the crew members´ reports are the observable phenomena in an emergency, and that the captain must control these. The results suggest that the best performing captain is shortly engaged in many groups on and offboard. Implications for design and evaluation of support systems are discussed.

Key words: End users, Evaluation, Man/machine interaction, Management systems, Observers.

1. INTRODUCTION

This paper presents observations from an emergency management exercise made as part of a three-day course in advanced fire figthing for ship officers. The exercise was based upon a simulation of a fire on board a merchant ship. In an emergency such as fire on board the participants take on different observer positions. Analytically at least three observer positions can be distinguished[1]. First, there is the traditional captain with the legally correct view of the captain during an emergency as the one with the overruling responsibility for the situation. This view focusses on the responsible captain´s will to stand back and, on basis of the information available on the bridge, deduce the solution to any problem that occur. Second, a more modern position is the shore based emergency response expert offering advice on a commercial basis. He must rely solely on the transmitted data and tentatively propose his conceptualisation of the emergency, which the captain may then accept or reject. Third and final, perhaps the most visible position is the fire party leader, who concentrate his effort on figthing the fire. There is a random element in his work, because he creates a fire party entrance with an eye on when and where in the ship a fire pops up. Now it should be noted that it is possible and in some cases even desirable for the captain to function as the expert or the fire party leader, e.g. when the ship is small or the crew restricted in numbers, then it could be appropriate for the captain to take part in the manual extinguishing of the fire.

[1]A taxonomy of decuctive, inductive and creative thought covers the cognitive aspects of the positions, leaving out free association and formal calculation (Johnson Laird 1993). It goes without saying that the captain in an emergency do not rely on free associations or heavy calculation.

In some of the positions computer based systems are used to support or control the observations. This is most notably in the emergency response centre ashore, where the experts qualify their discussions with calculation of worst cases, but available systems do also include onboard emergency management systems (Cowlan, *et al.* 1994; Hobday, *et. al.*, 1993; Rodricks, 1992). One of the purposes with the study presented in this paper was to contribute to the development and evaluation of a computer based support system of this type. The technology will be feasible on a ship equipped with a large number of sensors, a common data network and an integrated ship control centre. In traditional ships no computers are used for emergency management, though. The captain basically has four kind of media available: the layout of the ship, a list of crewmembers and their tasks, walkie talkies and satellite communication equipment, and blank sheets for recording actions taken. Other information is manipulated on top of these media. For example show threedimensional and deck by deck drawings the layout of the ship. Marks of smoke, fire and water are put on these drawings. Some areas of the ship, for example the accomodation area, are zoomed in on for further details of accessways/escape routes. A list of safety equipment on board presents the whereabouts of the hardware for fire fighting. The captain or an assisting officer ticks the mustered crew members on a muster list with a pen. Small red alarm lamps indicate activated fire detection sensors. Typically two assisting officers on the bridge, e.g. the third mate and the radioofficer, initiate and respond to walkie talkie and intercom communication with the fire parties and the engine room. The assisting officers furthermore timestamp the actions done by jotting down notes on a sheet of paper.

Such a use of media and assistants, as described in the above paragraphs, is quite common, and besides in cargo ships it is found in ferries and small passengerships. Now conflicting objectives of reducing crew size and increasing safety by introducing new technology makes it desirable to describe current work practice on board and how it can be supported (Lee and Sandquist, 1993). Designers ask how the captain is a necessary part of the emergency response, what he contributes with, how he interacts with others and what the essential features of his style of work are in case of an emergency?

Over the years a number of aspects of human cognition in the operation of ships has been studied, for example motor control in ship handling (Schuffel, 1986), time-distance perception in collision avoidance (Habberley and Taylor, 1989) and complex decision making in ship operation (Kerstholt, *et al.*, 1993), organisation of team navigation (Huthchins, 1988) and the learning of ship manoeuvring (Hansen and Clemmensen, 1993). A common theme in these studies is a view of the captain as an expert craftsman, who works skilfully within his familiar tasks. In this study of emergency management, however, the captain's work is peripheral and situation specific. His work is peripheral to the mate in charge of fire figthing and situation specific to the expert located ashore. An analogy to a very different work situation clarifies the attributes of the captain´s position. The Danish psychologist Moustgaard, who for many years have worked with psychological observation and description, refer to the classic demands to scientific observation: 1) the observer must be able to decide for himself when the process to be observed commences, 2) in a state of strained attention, the observer must study the phenomena and follow the course of their development, 3) for the results to be regarded as reliable, it must be possible to repeat each observation several times under the same conditions, and 4) the conditions under which the phenomena occur must be discovered by varying the circumstances and when discovered they must be systematically altered (Moustgaard 1992). These demands are interesting as norms for an ideal observer. The analogy makes it possible to question whether the captain during an emergency such as fire on board is in a similar classic, scientific observer position. The answer provided by this study was that, even though the captain only participates peripherally in the emergency, he is very active influencing the events.

2. METHOD

The empirical objects for the study of observation in maritime emergency management were a captain, two officers and manuals and plans on the fire figthing schools´ bridge training simulator, located at the Maritime Operations Center (MOC) at Southampton Institute, Warsash, UK[2]. The officers came from general cargo ships and channel ferry companies, and all had senior experience. Over a period of five months the exercise was repeated four times with different captains and crews. The exercise was based on paper and pen technology, with a few handbooks, plans and manuals available. The means for simulating the communication were walkie talkies and intercom for verbal interaction, and smoke indicators for signalling detected and acknowledged smoke. Instructors acted as the communication media to other ships and shore. The exercise was a "find the fire" exercise, where the participants were expected to react to smoke warnings not directly indicating the location of the fire. The part of the observed exercise was the participating captains´ and fellow officers emergency management on the bridge

[2]John Abell, chief instructor at the firefighting school at MOC, did select an interesting exercise for observation, and supported the study in other ways.

simulator. The exercise included some participants working in a connected engine room simulation and others working with a 2m high transparent plastic model of the ship with the location of the fire. A member of the staff at MOC videotaped the three bridge officers working, mostly focusing on the whole picture, but when he felt it was interesting, he zoomed in on the officers pen strokes on the plans etc. As an observer in all four exercises (and as a participant in one full course), this author had the opportunity to write notes when necessary, to discuss with the instructors during the exercise, and to walk around to get the different viewpoints from the other participants in the exercise. After each exercise, the instructors ensured that all participants had the oppurtunity to present their view of the exercise.

It was intended that the individual captain´s activity during fire on board and related damage stability problems should be studied. The captain´s activities have been defined in international recommendations on the safe operation of ships as primarily a matter of control[3]. The recommended captain behaviour in an emergency is to be motorically passive, to stay silent if possible, listen to the communication and monitor the instruments on the bridge. Only in some cases the captain needs to interrupt and control the operations in detail. Furthermore the intended objectfield was emergency management in high tech ships, operated on a one-man-bridge-operation concept, or on a pilot-copilot principle. This kind of operation are used on a number of Danish cargo ships and on some modern ferries, and to a certain extent duplicated from the operation of aircrafts. Generic for these principles for operation is that their function rely heavily on communication with crewmembers and passengers elsewhere in the ship, and on communication with helpers on ships nearby or ashore.

Theoretical reconstructions of high-tech operational schemes focus on individual captain´s function in military terms such as command and control or, more generally, as thinking machines (Johnson Laird 1993). A functionalist approach to the analysis of the emergency management activity points to a hierarchy of officers´ functions. The officer at the top level is the initiator of communication sequences, the person who acknowledges damage and identifies danger areas, who evaluates the consequence of further spread of water and fire and the person who decides when to stop the pumps and abandoning ship. Also the functional analysis of emergency management indicates that these functions will be activated by specific, recurring events during the emergency such as reports on fire, spread of smoke, missing crewmembers and hot spots on compartment divisions.

The results from a functional analysis should very much represent the bridge teams view of an emergency on board; but obviously other groups take part in what happens on board the ship during an emergency. Maritime authorities, such as government agencies, port authorities, coast guards, call or require calling in case of emergency. Classification societies offer satellite transmissioned expert advice. Ship companies have their own groups of experts, which are ready to be put on board. On board the ship the engineers are providing machinery power for steering and fire fighting purposes, while the purser and his employees take care of evacuation of passengers (if any). The mates work on the boat preparation party with communication equipment and supplies, and the fire parties use smoke diving equipment and fire hoses in their attempt to rescue people and contain the fire, eventually batten it down. Most important to the captain, the assistant officers work together with him on the bridge on plans and layouts. Now, the captain acts as the responsible person to all of these groups. The reconstructed objectfield of the present study in emergency management touched upon the ways the captain participated in each of these groups, in his unique position for observation in maritime emergency management.

3. RESULTS

The exercises at the fire fighting school were clearly influenced by psychological ideas of bottlenecks in the communication. Such ideas were probably an adequate conception of the way the participants had to perform in the bridge simulator with the present technology. Even the name of the course: command and control, communication, stage III and senior officers course seemed to reflect a view of emergency management as a hierarchical structure. Table 1 shows how the captain acted as the "bottleneck" in the flow of information. However, in the debriefing session afterwards, the captain and other participants stated that they came from ships equipped with better communications technology (more walkie talkies!), and they were not used to bottlenecks in the communication.

The firefigthing school was part of a large manufacturer of simulator training courses in the maritime field, and

[3]International conventions from the IMO (International Maritime Organisation) state that the master has as his responsibility that appropriate watchkeeping arrangements are made. The master and other key officers, who may have to control firefighting operations, should preferably receive advanced training in a number of items, that include fire control aboard ships. The more detailed standards for behaviour are then only visible in the available training courses in advanced firefigthing. This paper build upon observations made during such advanced fire control training.

Table 1 From-To communication pattern in the fastest performing captain and crew of four crews. In the uttermost left column the initiator of an utterance is listed and across the row the next speaker is listed.

	Crew					
From/ To	Ca	RO	3rd	Fp1	Fp2	ER
Captain (Ca)	-	19	16	44	15	10
Radioofficer (RO)	9	-	-	-	-	19
3rd mate (3rd)	26	-	-	-	-	-
Fireparty no 1 (Fp1)	42	-	-	-	-	-
Fireparty no 2 (Fp2)	12	-	-	-	-	-
Engine room (ER)	9	32	-	-	-	-

the site had a great diversity of simulators. The fire figthing command and control course were offered as part of a packet of training courses, so that different aspects of maritime work could be trained during one stay. Despite the fact that the chief instructor was a navigator, and the command and control exercise was designed for navigators, there was only little integration of manouevring on one hand and command and control training on the other. Had the exercise been performed in a manouvring simulator, the captain had probably worked with engine and rudder as well, in order to avoid the ship being filled with smoke. The captains did observe the wind direction, but in general they concentrated on information relevant for their men's immediate safety. Table 2 shows that most of the observations during the exercise had to do with the crewmembers safety. For example during these first fifteen minutes 81 observations (average of the four crews) were made regarding who were present and correct according to the muster list. Many observations (55, average of the four crews) were merely requests, accepts and initiation of situation reports, such as "Fireparty one?" or "Okay, Fireparty one" in the walkie talkie (WT) or the internal telephone (Intercom). Other frequent observations had to do with the location of crewmembers and how they should gain access to or escape from the area with the seat of the fire (28 and 27 observations respectively, as an average of the four crews). Much less frequent were the observations related to the ships survivability and to the operation of equipment, for example those observations concerned with the fire boundary plan and the fixed equipment plan (respectively 4 and 2 observations on average).

Observation results similar to those described in the above paragraphs were presented for the officers in the last of the four courses. Following this presentation the officers showed interest towards a comparision of performance across crews. Evaluation of ship officers was not a standard part of the fire figting courses, but according to the chief instructor it could easily be. Some of the good effects of evaluation, such as that caused by the present study, were that the participants worked harder and that they went for what they believed was a safe approach. One of the captains stated: "..but in general, those who try to support us generate more work for us..".

Simulator instructors at MOC argued that there was a need for certificating officers on ground of competence showed, instead of examinations in a school context. Experience showed large discrepancies between what the officers stated about their behaviour when interrogated in more formal sessions and what they did when at the bridge of a ship at sea. Qualified officers chose deliberately to ignore rules and regulations while at sea, and perhaps more relevant for the study presented here, while they were in simulators for training. The technological assumption behind the study of observation in maritime emergency management was that a proper psychological study would do for a better man-machine interface on an on-board support system, developed at the Danish Maritime Institute (DMI)[4]. Several strategies for reducing the effect of human error were discussed as relevant for the design of the prototype system. One assumption was that improvement of the captains work ability reduced the effects of accidents at sea. Alternative

Table 2. Objects observed by four different crews during the first fifteen minutes of an emergency exercise.

Exercise	1	2	3	4
Musterlist	104	79	82	60
WT/Intercom	90	29	51	51
3-Dimensional plan	27	22	30	32
Access/escape routes	33	40	27	9
Ventilation system/doors	40	12	27	21
Wind direction	18	6	13	11
Electrical power system	13	5	14	13
Log book	7	12	18	7
Verbal/visual/audible alarm	12	1	14	12
Fire main layout	12	2	11	9
Safety equipment list	4	4	18	5
Deck plan	1	11	11	4
Alarm history	3	12	3	6
Cargo list	6	2	6	5
Compartment plan	4	4	1	6
Boundary plan	3	3	6	2
Fixed equipment plan	0	6	1	0

[4]At DMI the study was a part of a research project in what was earlier known as an artificial intelligence group.

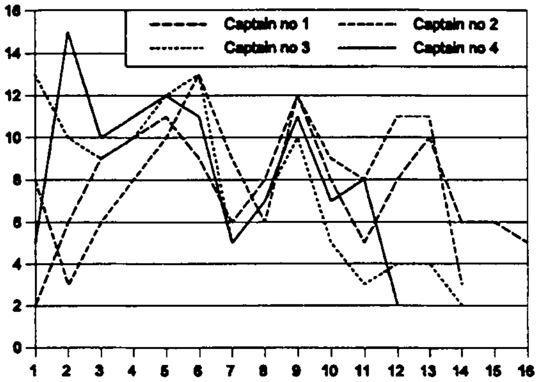


Fig. 1 The development of the captains initiatives per minute, counted as the number of utterances, from the missing motormen in the beginning of the exercise, and until they were recovered again. The captains´uniform reactions to significant information given by the instructors can be seen as peaks in the 6th and 9th minute of the exercise.

positions were expressed by officers in the debriefing exercises - that they have no influence on the course of accidents, which merely happen. A third alternative was that the officer should either accept or reject the support systems advice based on calculations of worst cases of fire and smoke spread. Figure 1 shows that perhaps all the strategies were relevant. In the first three minutes of the exercise the captains did differ in their approach, following information on two missing motormen. Then all the captains responded with an increase in observable activity towards advices on smoke in the alleway and hot spot on a bulkhead, delivered by the instructors according to a predefined scenario in the 6th and 9th minute. One of the captains, number four, was more effective and allowed the instructors to end this phase of the exercise early; i.e. two minutes before number two and three captain, and four minutes before number one captain.

4. DISCUSSION

The research in technological support for emergency management depended on a crossdisciplinary approach[5]. The empirical part of the study of observation in maritime emergency management was embedded in a human factors research programme at MOC. The theoretical approach taken in this programme was a combination of supervisory control, situation awareness and maritime collision avoidance studies, which influenced the study of observation mainly through the holistic approach to the operation of a ship.

The point of departure in the study was the cognitive theories of human error, which represents a model approach to analysis of work, presumably with power to predict human behaviour. More general, cognitive taxonomies of thinking are the conceptual apparatus with which the task demands on the captain are modelled. In the study of observation in maritime emergency management these demands were modelled as the thinking behaviour by the participants in the exercise when they performed their tasks. In principle this model should be convertible to a computer simulation of the participants behaviour, and tested against other behavioural data than those used for calibrating the model. Such data may be provided by future studies utilizing the sensor simulator capabilities of the prototype emergency management system.

In the study of observation of maritime emergency management the model of human performance was not applied without considering how the model depended on the observer´s position and the peripheral participation in the emergency. The historical assumptions behind the methodology of "observing" the captain´s observation in maritime emergency management was that observation can contribute with data which ultimately explain the essence of the behaviour observed. Introspection, measurement of motor and verbal behaviour, and etnographic depictions of the context of behaviour are some of the key words for the observation and description technology. The conditions for obtaining the goal of explanation have, however, worsened proportionally with the essence of objects being formulated by help of an increasing complex technology, so that any observation today depend on knowledge of the technology in the environment of the observer.

Cognitive psychology has developed concepts of thinking which has a clearly technical flavour, such as serial and parallel information processing, limited channel capaticy and coding information. The concepts were perhaps the only ones available to the instructors and the officers in the exercise. Eysenck and Keyne (1990) describe how cognitive psychology has origins in world war II Human Factors work. If this line of thinking is continued cognitive psychology is essentially a militaristic psychology, designed to conquer and rule over men and machines. But merchant officers cannot be trained to act as machines, rather their cooperation in an emergency require, that they adapt to each others pecularities.

Perhaps the captain´s behaviour only can be understood by focusing on the internal structures of emergency

---

[5]Research support was provided by the EU's EURET program from the Danish Maritime Institute in the ATOMOS (Advanced Technology for Optimization of Manpower On Ships) project. John Koch Nielsen was technical supervisor.

management, rather than on the observable motoric responses. The theoretical conditions for cognitive psychology were partly the behavioristic development of a scientific psychology and its failure of explaining psychological phenomena adequately and partly a general changing view of science from the activity of recording facts of nature to that of falsifying theories about the existence of an object. In a similar way, cognitive psychology in the maritime field may take advantage of authorities´ and owners´ growing distrust in natural observations of objects such as an officer´s competence.

Observation and interview may be closely knit together for the captain, who reacts to significant information and does away with the rest. By allowing his crew to interact with him and in between themselves, he controls the observations of fire and smoke made under time pressure. Repeated use of the available media provides the captain with a continous, stable view of what the crew observes. This paper on the captain´s observation in maritime emergency management attempt´s to contribute to a kind of maritime psychology. Theoretically the captain can be considered an important individual in an emergency, though his positions for observation seems to be much more diffuse and changing than a rational account may consider.

5. PERSPECTIVE

Methodologically this study should promote the use of human factors metodologies and aspects in the maritime industry, eg. development of task scenarios, behavioural case studies, and communication analysis. Theoretically it might contribute to a domain specific model, which has some interesting components. A first most important activity should be to observe behaviour on different type of ship or routes in direct relation to the captain´s work in an emergency. Also shorebased communication should be taken into account, if future studies should avoid being critizised as ungrounded, not scientific responsible.

Perhaps the captains situation in an emergency is a little like being a pilot on board a captain´s ship, or being the captain of a new ship with a different bridge layout (see e.g. Schragen 1994 and Schuffel 1992).

6. REFERENCES

Cowlan, N.J., Clemmensen, T., Cooke, M.J., Dhanjal, S., Earthy, J., Edmondson, P., Johnstone, J.L., Munden, A., and Nielsen, J.K., (1994). *Emergency Management System, Evaluation Report*, Final Report ATOMOS Task 2304, Lloyds Register, London.

Eysenck, M.W. and Keane, M.T., (1990). *Cognitive Psychology*, Lawrence Erlbaum, UK.

Habberley, J.S. and Taylor, D.H. (1989). Simulated collision avoidance manouvres: a parametric study. *The Journal of Navigation.* Vol 42. No2. , p248-254.

Hansen, J.P. and Clemmensen, T., (1993). Cooperation and learning in simulated ship manouvring, in: *Designing for Simplicity, Proceedings of Fourth European Conference on Cognitive Science Approaches to Process Control* , August 25-27, Frederiksborg, Denmark.

Hobday, J., Bright, C., Earthy, J., Rhoden, D., and Jones, P, (1993). *Supporting emergency and damage control on-board ship: A case study of information retrieval within multi-media applications.*, IEEE Multimedia Symposium, Southampton, UK.

Hutchins, E., (1989). The technology of Team Navigation. In: Galegher, J., Kraut, R. and Egido, C. (Eds.) *Intellectual Teamwork: Social and Technical Bases of Cooperative* Work. Hillsdale, NJ: Lawrence Erlbaum Associates.

Johnsohn-Laird, P.N., (1993) *Human and machine thinking*. Lawrence Erlbaum, London.

Kerstholt, J.H., Passenier, P.O. and Schuffel, H., (1994). *The effect of a priori probability and complexity on decision making in ship operation.* TNO-TM 1994 C-22, The Netherlands.

Lee, J.D. and Sanquist, T.F., (1993). A systematic evaluation of the impact of technological innovation on ship navigation, in: *The Impact of New Technology on the Marine Industries*, Warsash, Southampton Institute, UK.

Moustgaard, I.K. , (1990). *Psychological observation and description.* Sigma Forlag, Bergen.

Rodricks, B.S., (1993). *Intelligent fire and safety management systems for passenger vessels.* Cruise & Ferry 93 Conference papers & Speakers profiles, London.

Schraagen, J.M., (1993). *What information do river pilots use?*, TNO, The Netherlands.

Schuffel, H., (1986). *Human Control of Ships in Tracking Tasks.* Ph.D. thesis, TNO, The Netherlands.

LOOP-SHAPING CHARACTERISTICS OF A HUMAN OPERATOR IN A COMPENSATORY MANUAL CONTROL SYSTEM

Takeshi INABA and Yoshiki MATSUO

*Dept. of Control and Systems Engineering, Faculty of Engineering,
Tokyo Institute of Technology,
2-12-1, O-okayama, Meguro-ku, Tokyo, 152, JAPAN*

Abstract: New observations on loop-shaping characteristics of a human operator in a compensatory manual control system are presented. These observations are quite useful for manual control systems design. A design method of manual control systems considering easiness of the operation is proposed based on H^{∞} control theory. The validity of the design method is confirmed by experiments.

Keywords: Manual control, Manual operation, Man/machine systems, Human-machine interface, H^{∞} control

1. INTRODUCTION

This paper deals with new observations on loop-shaping characteristics of a human operator in a compensatory manual control system and their application to a design method of human-oriented manual control systems considering operator's easiness.

In the late 1950s to 1960s, studies were presented on analysis of human operator characteristics in compensatory manual control systems, based on the classical feedback control theory in the frequency domain. But their results were concerned only with properties near the gain crossover frequency. Besides, there have been few studies on application of them to manual control system design.

In this paper, these classical results are expanded in order to apply them to design. At first, we performed experiments on a compensatory manual control system with the control object realized on a computer. The experimental results are analyzed carefully not only at frequencies near the gain crossover but at the lower frequencies. As a result, new observations including loop-shaping characteristics of a human operator are obtained.

Secondly, using these observations a new design method of manual control systems with regard to operator's easiness based on H^{∞} control theory is proposed.

Finally, validity of the design method is confirmed by experiments.

2. COMPENSATORY MANUAL CONTROL SYSTEM AND CLASSICAL RESULTS

A compensatory manual control system is illustrated in Fig.1. In the late 1950s to 1960s, several studies were presented on analysis of human operator characteristics in that system, based on the classical feedback control theory in the frequency domain. Their main results are as follows (McRuer, *et al*, 1967).

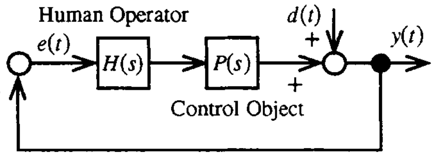


Fig.1. Block diagram of a compensatory manual control system.

a) A human operator in a compensatory manual control system adjusts his transfer characteristics to his control object. And the class of possible transfer functions of the human operator after the self-adjustment is clarified as following transfer function model.

$$H(s) = K\frac{T_L s + 1}{T_I s + 1}e^{-\tau s} \qquad (1)$$

b) After the self-adjustment, the open loop transfer function of the compensatory manual control system $H(s)P(s)$ can be approximated to a transfer function (equ.(2)) at frequencies near the gain crossover, which is called the crossover model.

$$L_{CO}(s) \equiv \frac{\omega_c e^{-\tau s}}{s} \qquad (2)$$

Explanation of the above properties of a operator in view of the classical feedback control theory is as follows. In a compensatory manual control system the operator tries to shape the open loop transfer function to the above crossover model (equ.(2)) at frequencies near the gain crossover by adjusting his transfer characteristics within the model (equ.(1)). Thereby, he keeps some amount of stability margin of the system.

But these models are not enough for designing the manual control system because of the lack of knowledge on the properties of the operator at low frequencies which are related to the closed loop performances such as disturbance rejection. Accordingly, there have been few studies on application of these models to manual control system design.

3. LOOP-SHAPING CHARACTERISTICS OF A OPERATOR AND OTHER NEW OBSERVATIONS

3.1 Configuration of the measurement system

Fig.2 illustrates experimental equipment used for the measurement of a human operator's transfer function in compensatory manual control systems. The control object and random disturbances are realized on a computer. And a human operator tries to eliminate the output error signal indicated on a CRT display using a rotational input device. About sixty types of stable transfer functions are used as the control object.

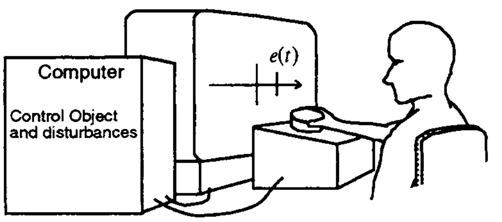


Fig.2. Experimental equipment for the measurement.

3.2 Confirmation of the classical results

Fig.3 shows the open loop transfer characteristics of the system for various control object. These results show that a operator seems to adjust his transfer characteristics to the control object within the model (equ.(1)) and all the open loop transfer characteristics are close to the crossover model (equ.(2)) near the gain crossover frequency. This confirms the classical results mentioned in the previous section.

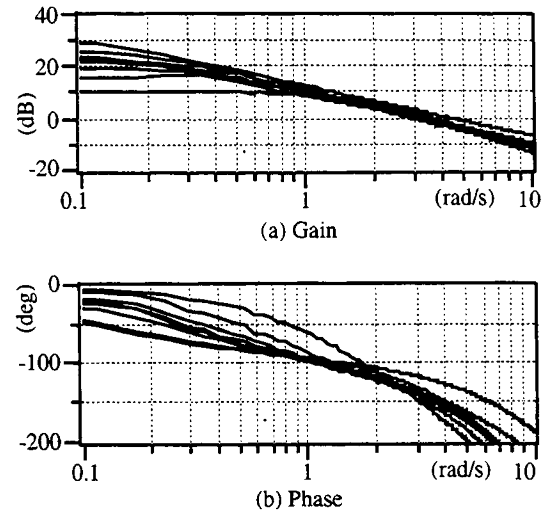
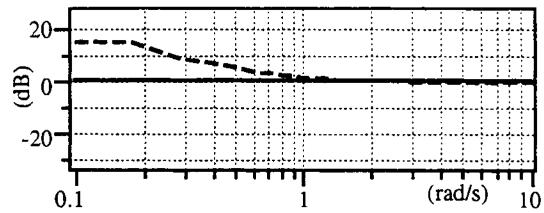


Fig.3. Open loop transfer characteristics for various control objects.

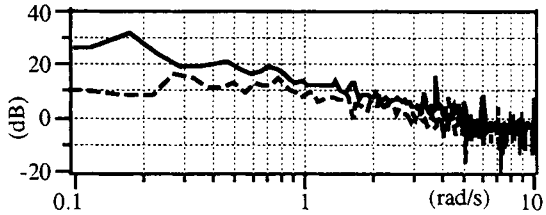
3.3 Loop-shaping characteristics of a operator

In order to obtain new observations on properties of the operator at low frequencies, which the classical results lack, precise analyses are carried out. Experimental results for two control objects which have different characteristics at low frequencies are showed in Fig.4. As the dashed line in Fig.4(b) indicates, the operator adjusts his transfer characteristics at low frequencies when the control object increases low-frequency gain as the dashed line in Fig.4(a) shows. The adjustment is in such a way that two open loop transfer characteristics are close to each other even at low frequencies as showed in Fig.4(c). This means that the operator operates to shape the open loop transfer characteristics into a transfer characteristics represented by equ.(3), that is to say, he achieves a *loop-shaping*. Let us call equ.(3) the loop-shaping model.

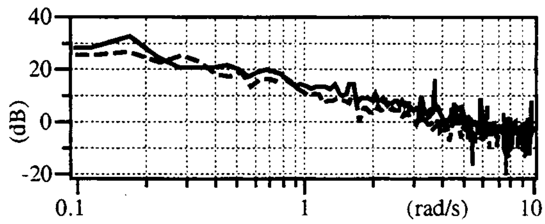
$$L_{OL}(s) \equiv \frac{T\omega_c e^{-\tau s}}{Ts + 1} \qquad (3)$$



(a) Control Object Gain



(b) Operator Gain



(c) Open Loop Gain

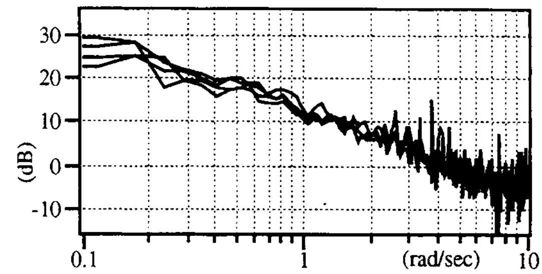
Fig.4. Loop-shaping characteristics.

3.4 Relations between operator's transfer characteristics and easiness of operation

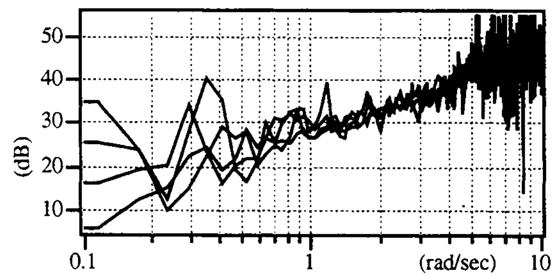
On designing manual control systems, regard for the operator's of operation is significant. Therefore relation between operator's transfer characteristics and the easiness is explored. Fig.5 shows experimental results of operator's transfer characteristics for two control objects, for one the operator behaves like integral element (Fig.5(a)) and for the other like differential element (Fig.5(b)). On each figure four curves correspond to four trials of the same operator. Fig.5(a) indicates that the operator shows almost the same integral-like characteristics at frequencies lower than the gain crossover frequency on each trial. In these cases, the operator feels easiest in operation. However, as in Fig.5(b) the similarity of the operator's transfer characteristics is not observed when himself behaves differential-like except near the gain crossover frequency. In these cases, the operator seems to concentrate on near the gain crossover frequency to stabilize the system and cannot pay attention to operations at low frequencies.

3.5 Performance of a manual control system.

A compensatory control system should have good disturbance rejection, that is, small sensitivity of the system at low frequency. Fig.6 shows experimental results of the closed-loop sensitivity characteristics for the two control objects. Better sensitivity characteristics (namely, less sensitive) at all frequencies lower than the gain crossover frequency is obtained when the operator shows integral-like characteristics (dashed) than deferential-like one (solid). Moreover, in various experiments, it is observed that when the operator achieves the loop-shaping mentioned before, the system has the best sensitivity characteristics.



(a) Integral-like



(b) Differential-like

Fig.5. Operator's transfer characteristics.

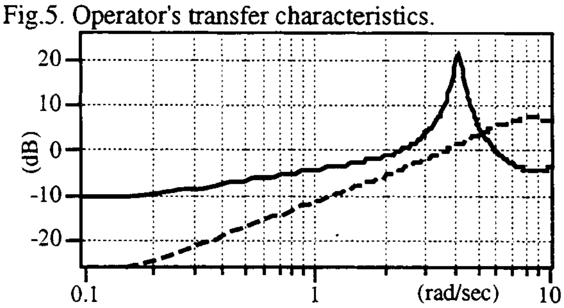


Fig.6. Closed loop sensitivity characteristics.

New observations stated in this section are summarized as follows.

New observation 1: The loop-shaping characteristics; A human operator operates to shape the open loop transfer function into the loop-shaping model (equ.(3)) at all frequencies lower than the gain crossover frequency when the operator shows integral-like characteristics near the gain crossover frequency.

New observation 2: The best easiness of the operation; A human operator feels easiest when himself behaves integral-like.

New observation 3: The best disturbance rejection; A compensatory manual control system has the best sensitivity characteristics when a operator shows integral-like characteristics and he achieves the loop-shaping.

4. SYNTHESIS OF MANUAL CONTROL SYSTEM AND EXPERIMENTS

Using proposed new observations it is possible to synthesis manual control systems considering not only the closed loop stability but also the disturbance rejection and the easiness of the operation.

4.1 Configuration of the manual control system with dynamic compensators

Fig.7 illustrates the considered manual control system. Dynamic compensators $C_1(s)$, $C_2(s)$ are inserted between the human operator and the control object. L.P.F. is a low pass filter. In a compensatory manual control system, presentation of unnecessary high frequency components of the feedback signal makes the human operator decrease the crossover frequency (McRuer, *et al*, 1967). So the filter is inserted.

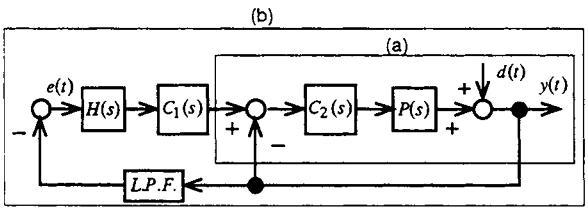


Fig.7 Configuration of a manual control system with dynamic compensators.

4.2 Design procedure of the dynamic compensators

At first, the dynamic compensator $C_2(s)$ is designed so that the system without the operator (illustrated as (a) in Fig.7) rejects the disturbance sufficiently. The compensator is designed by applying H^∞ control theory to an augmented system including frequency dependent weighting functions $W_1(s)$, $W_2(s)$ (Fig.8). It is assured that the obtained compensator internally stabilizes the system and bounds the H^∞ norm of the closed loop transfer function matrix from the disturbance v to the controlled output z below the specified value. According as the gain of the weighting function $W_1(s)$ is increased, the gain of the closed loop sensitivity function is decreased. The weighting function $W_2(s)$ is used for specifying robust stability.

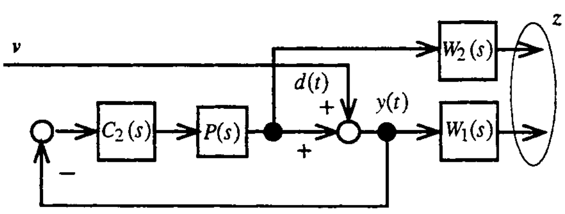


Fig.8. Augmented system for the design of $C_2(s)$.

Next, the dynamic compensator $C_1(s)$ is designed using new observations. Fig.9 illustrates the considered augmented system. Here the compensator $C_2(s)$ is fixed as designed before. From new observations 2 and 3, the best easiness of the operation and the best disturbance rejection is obtained simultaneously when a operator behaves integral-like. Therefore a transfer function with this characteristics is used as the operator model $H(s)$. However, the operator does not necessarily behave like this model. From new observation 1, the operator tries to achieve the loop-shaping, if possible. So, if the open loop transfer function of the considered system is close to the open loop model (equ(3)), the operator achieves the loop-shaping and his transfer characteristics must be close to the model $H(s)$. The compensator $C_1(s)$ is also obtained using H^∞ control theory using by the ability to shape transfer functions with frequency dependent weightings, $W_3(s)$ and $W_4(s)$. The open loop transfer function with the obtained controller is close to $W_3(s)$ at the frequencies where $|W_3(j\omega)| >> 1$ (typically at low frequencies) and is close to $W_4(s)^{-1}$ at the frequencies where $|W_4(j\omega)| >> 1$ (typically at high frequencies). From this fact, the open loop transfer function of the system can be specified by these weights.

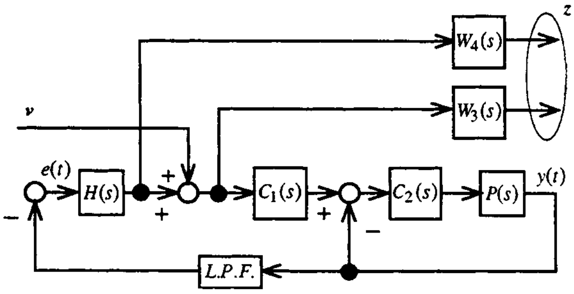


Fig.9. Augmented system for the design of $C_1(s)$.

4.3 Design example and experimental results

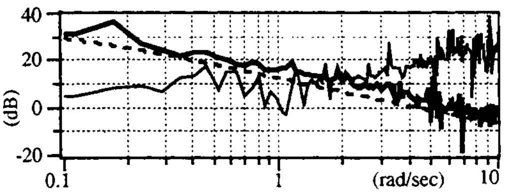
The synthesis and experiments were carried out for a low damping second order lag element as the control object which is very difficult to be controlled without the compensators. All transfer functions used in this example are listed on Table 1.

Fig.10 shows experimental results with the designed dynamic compensators. Fig.10(a) shows that the operator's transfer characteristics (heavy) is close to the model $H(s)$ (dashed). And from Fig.10(b) it is confirmed that the open loop transfer characteristics of the system with the compensators (heavy) is also close to the loop-shaping model assumed in the synthesis (dashed), which implies the operator achieves the loop-shaping. Fig.10(c) shows that great improvements on the disturbance rejection properties are obtained with the compensators.

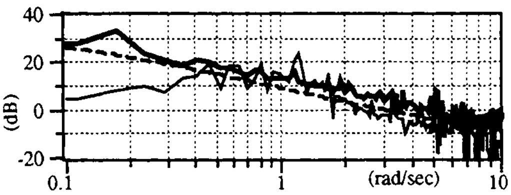
These results verify the validity of the proposed synthesis method.

Table1 Transfer functions used in the design example.

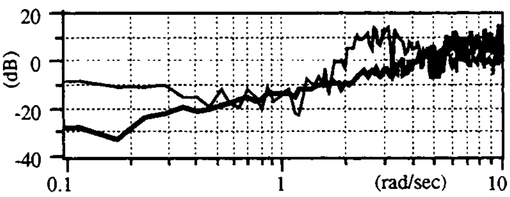
$H(s)$	$\frac{2.1s+42}{10s+1}e^{-0.2s}$	$W_1(s)$	$\frac{180}{10s+1}$
$P(s)$	$\frac{1}{s^2+0.2s+1}$	$W_2(s)$	$0.01(s+1)^2$
$L_{OL}(s)$	$\frac{30}{10s+1}e^{-0.2s}$	$W_3(s)$	$\frac{29}{10s+1}$
$L.P.F.$	$\frac{1}{(s+1)^2}$	$W_4(s)$	$\frac{0.1s+0.001}{0.017s+1}$



(a) Operator Gain



(b) Open Loop Gain



(b) Sensitivity Function Gain

Fig.10. Experimental results with the designed dynamic compensators.

5. CONCLUSIONS

In this paper, new observations on the properties of human operator in a compensatory manual control system were clarified. And a design method of human-oriented manual control systems using these observations was proposed.

ACKNOWLEDGMENT

This work is supported by Tateisi Science and Technology Foundation.

REFERENCES

McRuer, D. T., Jex, H. R. (1967). A Review of Quasi-Linear Pilot Models. *IEEE TRANSACTION OF HUMAN FACTORS IN ELECTRONICS*, **HFE-8 No.3**, pp. 231-249.

HOW DO INDUSTRY DESIGN ASSEMBLY SYSTEMS A CASE STUDY

M.Sc. Monica Bellgran*, Prod. Eng. Erkki Lahti**

**former Lundström, Div. of Assembly Technology, Dept. of Mechanical Engineering Linköping University, S-581 83 Linköping, Sweden.*
*** FHP Elmotor AB, S-590 90 Ankarsrum, Sweden.*

Abstract: This paper presents an ongoing case study where a developed method for systematic design is applied when designing an assembly system at a Swedish company, FHP Elmotor AB. When initiating the method, it became obvious that the initial system background was not sufficiently analysed. Therefore, a design team was formed and a thorough background analysis was performed. The original goal description had to be changed as well as the product design. Using the design method made the company's internal system design process more important than initialy intended.

Keywords: Assembly, systems design, systems methodology, industrial production systems, project management.

1. INTRODUCTION

Traditionally, the process when designing manufacturing and assembly systems is not emphasised to the same extent as the process of designing products. This situation concerns both the industrial companies' focus, and the scientific research made in the field of engineering and manufacturing design. There exist for example relatively few methods and principles developed in order to improve the assembly system design process. However, a product of high quality - referring to the total quality concept - cannot be assembled without a suitable assembly system which in turn requires an emphasis on the design process.

One way of improving the assembly system design process is to structure and systematise the process and to emphasise essential factors affecting the system and its design. For this purpose, a first version of a structured method for systematic design of assembly systems has been developed at the division of Assembly Technology at Linköping University (Lundström, 1993a).

1.1 The Background of the Development of a new Assembly System using the Design Method

FHP Elmotor AB, a Swedish company of 450 employees located in Ankarsrum, Sweden, was going to design a new assembly system for a new product generation. The developed system design method was implemented for the purpose in early 1994.

Since 1990, FHP Elmotor AB belongs to the FHP Motors GmbH group, one of the major producers of electric motors in Europe, 50 % owned by Electrolux and 50 % owned by Temic (a Daimler-Benz subsidiary). FHP manufactures six different product families, and among these, the specific product which is discussed in this paper. A new product generation of the product is being developed, and the manufacturing of this new product is planned to be initiated in a few years. The product, the annual production volumes, year and investment costs are confidential due to the current development stage of the project.

The existing assembly system which was built during the eighties, consists of three subsystems and the final assembly system. The subsystems consist of dedicated machines combined in a line flow, resulting in a rather inflexible system with high set-up times, and buffers between the stations. The final assembly is completely automated, performed in one machine which requires supervision.

The features of the new product are such that it cannot be assembled in the existing system. One reason is that the product design and structure are different. Another important reason is that the equipment of the existing system is old and needs to be updated. A new system must therefore be developed for the new product. The question is: *how to design this new assembly system in order to achieve a suitable system?*

The manufacturing and the assembly process of the product is essential for the company. Customer adapted products as well as the customers' requirements of short lead times etc. require a flexible production process. The new assembly system is therefore of significant importance concerning market aspects.

The objective of the case study is to answer the question of how this specific assembly system is designed at FHP when using the design method. Who designs the system and what the result of the design process is, are other related questions to be answered in the case study. Another important question is how internal and external factors may affect the preconditions of the design process.

This paper presents the results achieved so far, of how the company designs its assembly system when using the method. Specific factors at the company affecting the design process are indicated, and aspects of the method are discussed in the paper. The case study is planned to continue until the assembly system is built and is running in full production.

2. THE METHOD OF RESEARCH

The case study at FHP Elmotor AB is an opportunity to study and verify the application of the developed method for design of assembly systems. The application of the design method was initially introduced in February 1994 by the project leader after studying the design method. The applied system design method is a theoretical foundation mainly based on a thorough analysis of the problem area and its related research, but also based on studies of assembly system activities in industrial companies.

This research project is performed by the utilization of qualitative research methods. The study at FHP suits the definition of a case study as an empirical inquiry according to Yin's definition (Yin, 1981). The research questions of *how* assembly systems are designed, and *why*, also fit into the general case study strategy. A single-case design is used, involving multiple units of analysis in order to study how the company designs the system.

The research method may partly be considered as a type of case study called action research, where the researcher participates actively in the change process. In this case, the research participation mainly concerns the transfer of knowledge through the initial introduction of the system design method.

Multiple sources of evidence are used, which is a major strength of case study data collection according to Yin (1981). Continuous follow-up of the project status, by contacts with for example the project leader, is one technique used. Other techniques are structured questionnaires and documentation analysis. Interviews with members of the project group and other employees will follow at a later project stage. The contextual issues are important in this case and a broad analysis of the company is therefore necessary.

3. THE SYSTEM DESIGN PROCESS

Two important and closely related aspects to focus on when designing the assembly systems are the design process and the designers. As concerns the design process, it is often based on company's traditions, the experiences of the designers or the advice of external consultants. Another solution is to have suppliers of assembly equipment to design and/or to build the systems. In general, the design process is to a large extent determined by the designers.

Systematic design principles or methods are seldom used by industry when designing assembly system. However, when designing production systems for example, the establishment of a set of manufacturing policies has been done according to Skinner (1978). Related work in the assembly system design field, however with certain different approaches, can be found in for example Milberg (1989), Maczka (1985), Warnecke (1987) and Feldmann (1990).

3.1 Method Presentation

In order to improve the system design process, a first version of a systematic method for the design of assembly systems has been developed. Here, the design process is sequentially structured and relevant parameters are identified and classified into categories (Lundström, 1993a).

The structure of the design process involves three blocks; the background, the development and the realization of the assembly system, see figure 1. These blocks comprise thirteen phases in total, which are further sub-divided into design steps on four levels of detail in the form of a tree structure. This structure makes it possible for the designer to choose level of the support needed. The presentation principle is based on IDEF (Wu, 1992). Some general principles of importance to the method are that:

- *The designers' ability to solve the problem of designing the system should be supported.*
- *Automatic or manual assembly systems, or a combination of these systems are considered.*
- *The technical system and the social system (the work organization and the work environment) is equally important.*
- *The background analysis is emphasized in order to gather relevant information before developing alternative system concepts.*

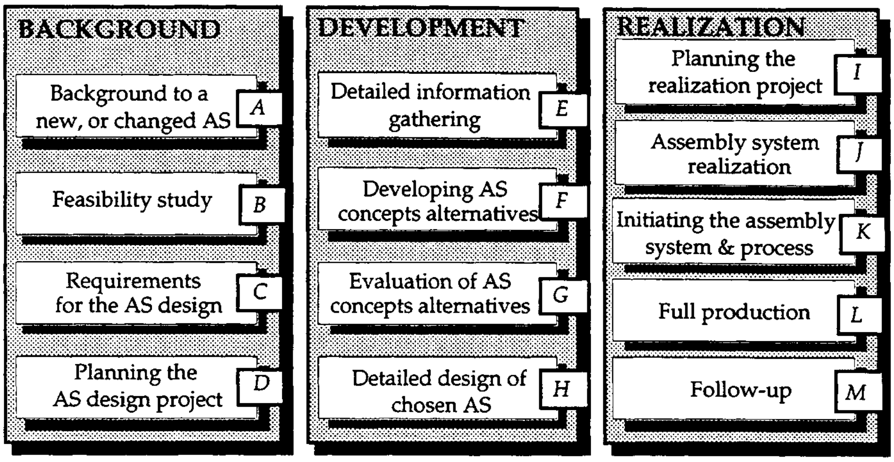


Fig. 1. The system design method comprising the 13 phases and structured in the three blocks; background, design and realization. (The figure is slightly modified from Lundström and Johansson (1993b).

3.2 Applying the System Design method at FHP

The development of the new generation of the product began in 1993. In late 1993, the work of how to design the new assembly system for the product was initiated at the production engineering department. A supplier of automatic assembly equipment was contacted at that early stage of the design process and was asked to submit an offer. A principle layout was developed, see figure 2, and the number of operators working in the system were suggested as well.

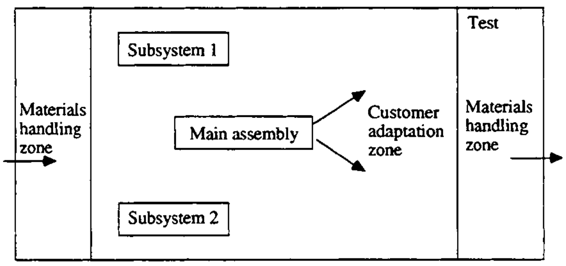


Fig. 2. The first preliminary layout of the assembly system at FHP Elmotor AB.

In the beginning of 1994, a new production engineer was employed. He become the project leader of the system design project, and initiated the implementation of the design method at this project stage.

After initiating the method, it became obvious that the design process had passed fairly quickly through the different design phases. As a result, the background of the project was not sufficiently analysed. Detailed solutions had been suggested and a supplier had been appointed too early in the design process.

This situation is not unusual. Often, when the economical resources are determined, the best technical system alternative (or supplier) is chosen, and then an adaptation to the social system is made. This often results in problems in the initial production stage.

Traditionally, assembly systems at FHP have often been designed based on the selection of the best technical assembly system concept, developed by suppliers of assembly equipment. The time for the design process has often been short, and the internal responsibility has mainly been put on production engineers alone instead of appointing a team of individuals with different skills.

Assembly systems in general are often designed by production engineers within the company, or by a team composed of people with different skills. The system design process at FHP so far, proceeded by developing the project team in three steps:

Step I. The initial system design work was made by production engineers at the Production Engineering Department.
• No system design method was used.

Step II. In the next step of the process, a production engineering manager, a representative from the economics department and the newly employed production engineer formed the system design team. The production engineer became the project leader.
• The system design method was initially implemented.

Step III. When the project increased to comprise the building and the materials handling as well, a formal project team was initiated. This team concluded the managers of eight different departments and a control group (top management). The former project leader of the system design became responsible for the total project involving both product and system development.
• The system design method is implemented and used. This is the current project status.

The product was developed by a separate project team which had already been initiated in step I, when also the system design process was initiated. The separate product development process continued in parallel with the system design process (step II). In step III, the system design and the product development was integrated in the total project. The project team now comprises 9 persons; the project leader and the managers of the departments of Production, Manufacturing Planning, Quality, Research & Development, Marketing, Administration, ADP, and Purchase, see figure 3.

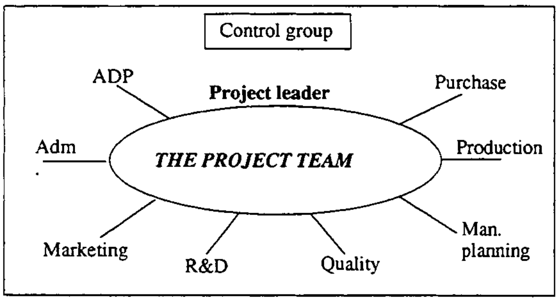


Fig. 3. The project team at FHP.

When designing the assembly system in a more systematic way using the method, the background to the task became important again and was carefully analysed at FHP. The project has proceeded so that phases A, B and C of the design method has been worked through on different levels of detail, see figure 1. It included an extensive analysis of the background of the necessity to design a new assembly system at FHP (phase A). The feasibility study investigating whether it was possible to perform the project referring to available resources etc. (phase B), and the important specification of the strategies and requirements of the potential assembly system (phase C).

More specifically, the system design process at FHP following the first method phases, involved for example the investment offers, the composition of the project team, and a thorough background analysis. This system design process proceeded in parallel with the product development process.

3.3 Focus on the Background Analysis

As a part of the background analysis when designing the assembly system, a problem analysis followed (method phase A, design step A2, see figure 4). Here, operators were working in two groups in order to identify problems mainly with the previous product and the existing assembly system, in order to gather information and specify requirements for the new system. The result of this analysis mainly concerned the assembly operations and problems related to equipment, but also opinions regarding a new assembly system concept and aspects on how to work in it.

When the design team (see step III) had been determined, a very important background analysis was made within the project team members. The objective of this problem analysis on a strategic level was to answer the questions; why?, what? and how? with respect to how each department was working. Certain design steps of phase A of the design method were selected, see figure 4.

Phase A: Background to a new, or changed assembly system.

A1: Analyse the preconditions
A11: Why change?
A12: Resource specification

A2: Problem analysis
A21: Problems in the existing organisation
A22: Reasons for the problems

A3: Bench-marking
A31: What is to be bench-marked?
A32: How to bench-mark?
A33: Specification

A4: Requirement analysis
A41: Define requirements classified as strategic system parameters
A42: Define requirements classified as operative system parameters

Fig. 4. The selected and briefly modified design steps from phase A of the method, used in the background analysis conducted by the project team.

The outcome of this analysis was important; problems and various aspects such as strategic issues, the work organisation, education demands etc. were focused. The connection between the different departments at the company such as Quality - Production, Quality - Purchase etc. was emphasized, which resulted in a mutual insight into what would be required, areas of responsibilities etc.

Another concrete result of this background analysis was that it became obvious that the original goal description of the total project had to be changed. This was mainly based on conventional ideas, and would probably have resulted in the development of an assembly system, similar to the existing one.

The first investment request of the project was based on the original background analysis, which in turn was partly based on the information received from the supplier concerning the proposed technical system. However, this information was not sufficient enough for the owners of the company to make the decision whether the project should be financed or not. Therefore, the financial decision was postponed until sufficient information was available. The thorough background analysis made when using the design method became important to achieve a sufficient foundation for the coming investment decision.

The postponed investment decision affected FHP's possibilities to continuing with the project at full speed. However, the original timetable for the project still remains due to the reduction of time for the realization of the system.

The system design method was introduced at FHP at a stage when a prototype of the product existed. However, after the thorough and more systematic background analysis it became obvious that the product needed to be developed further. A new concept was necessary where the manufacturability of the product was taken into account to a larger extent. The development of the product therefore still continues.

4. DISCUSSIONS

4.1 The Design Process and the Designers

When the system design project increased at FHP due to new requirements, it was necessary to initiate a team consisting of representatives from the different departments concerned at the company. It was also advantageous to involve individuals with the authority to make significant decisions. A team's success is determined by the integrated outcome of everyone's work. Development projects are typically conducted under intense time and budget pressures, and therefore, they usually magnify the strengths and weaknesses of a company (Bowen, 1994). Another aspect is the size of the project team (9 persons), all with different skills. According to Oliva, it is advantageous if a team is composed of a small number of people (8-15) where each member is a specialist in his own field (Oliva, 1990). The formal creation of the project team, and also the simultaneous design of the product and the assembly system, may give the conditions for a better understanding of these design processes.

The choice of project leader is important in any project. This system design project is no exception. Intrapreneuring is what may be described as "acting like an entrepreneur inside a large organisation" (Robson, 1990). Having intrapreneurs - or passionate champions - within a project is very important for the success of the project. Referring to a study of 50 new products made by Texas Instruments, Robson states that all projects that failed lacked a passionate champion.

In order to motivate the use of the method when designing, a visualisation of the project description concerning the background phases was made at FHP. The project was resembled with a vehicle going on a long trip. Here, the background studies would give answers to what type of vehicle that was used, possible problems of this vehicle, information of the best existing vehicle, general and specific requirements and the possibilities to fulfil these (see chapter 3.2).

Operator involvement is also important when designing. So far, operators at FHP have been involved in analysing the problems of the existing product and assembly system. Except the gathering of significant information, this operator involvement gives the preconditions for anchoring and later acceptance of solutions. This may increase their motivation for changes. The system design method further indicates that operator involvement is also important when specifying the assembly system requirements, and when developing assembly system alternatives at the conceptual level.

4.2 Factors Affecting the Design Project

The assembly system design does not always proceed as planned, due to both internal and external factors. At FHP, the preparation of information as a basis for the financial requirements, as well as waiting for the investment decision, had a controlling affect on the first phases of the project. One reason for this was shortage of information, needed to make the decision of whether financial support would be approved or not. In this context, the organisation and geographical location of the ownership may be of importance. The complexity of the communication channels may affect the decisions.

The project lead times may be affected if the investment decision is postponed. The consequence may be shortening of the important project planning stage. Another scenario is to continue with the system design while waiting for the financial decision, taking a risk that there will be no investment approval. A difficult aspect of the investment inquiries concerning the assembly system design is to make a realistic estimation of the costs. This estimation of the project cost can be made by experienced project participants. FHP requested a budget proposal from a supplier of automatic equipment, which indicated the size of the investments. This is another way of estimating the costs of the system. However, relying on the suppliers' proposal implies a risk of deciding too early what type of system concept to choose. After applying the design method, FHP has found it important to gather relevant information before developing any system alternatives, in order to develop the most suitable assembly system.

The selection of both internal and/or external designers is very important when designing an assembly system. An important principle of the method is to handle as much as possible of the design process within the company. Referring to the case study, the assembly system conditions are best known to FHP, and it is therefore advantageous if the important analysis and the specification for the system design are made by designers within the company. As a result, knowledge and experience regarding the system will be maintained within the company, which can be considered as "to own the technology". This may for example facilitate the knowledge-transfer to the operators at a later stage. However, the actual building of the system may be realised by a supplier on the basis of the requirements specified by the company. Whether external competence, such as suppliers are contacted and required for the system design, depends on the company's knowledge and ability to develop and build the system. The advantages of co-operating with suppliers are the utilisation of their knowledge and experience in the system design field. The risk, however, is the use of existing applications and modifications of old technical solutions that do not suit the new tasks. This risk may be greater if the feasibility study made by the supplier is not externally financed.

5. CONCLUSIONS

The process when designing the assembly system at FHP is based on the principles of the system design method. It has resulted in a thorough background analysis which has taken more time than initially planned. However, emphasizing the background and the feasibility study is here the means of reducing the time for the realization and the initiation of the new assembly system. Therefore, it is estimated that the original project timetable will still be maintained. Instead, focus has been moved to the initial stage of the system design process. By emphasizing the problems at an early project stage, the probabilities of late surprises have decreased, and the possibilities for higher system quality may increase.

By emphasizing on this early stage, the preconditions for co-operation over the department walls have increased as well. It has become more natural to work across the department boarders.

The newly employed project leader of the project was not influenced to the same extent as old employees by the traditions of the company of how to design assembly systems. This probably made the background analysis become more evidently important as both the product and the existing system was new. The use of the method at the company has matured the more working with it, and it has been found advantageous to work outgoing from a structure and being more systematic when designing the system.

Over a longer term, the advantages by designing systems systematically using a method, means having something to refer to when improving the design process. When outgoing from a structure, it is easier to find out where it went right or wrong, which makes it easier to adapt the design process to the next project. Developing a system for applying what was learned from one project to subsequent projects is an important key according to Bowen (1994).

When focusing on the design process, the possibility to achieve higher system quality, to reduce problems at an initial production stage and to minimize late system changes increases. The overall goal is to reduce the total lead times and costs for the development of a system.

The use of the method is based on working iteratively, choosing design steps in an order that is suitable for FHP. Applying the method has also resulted in the emphasis on certain important questions at issue. So far, the knowledge of the assembly system design is still maintained within the company.

According to the project leader at FHP, the system design method is quite easy to learn, but what may take time is changing the attitudes among people at the company so the necessity of being systematic when designing becomes accepted. An important task is to motivate employees at all levels for the project. This concerns both the project team members, the operators and the management. Working with the attitudes and the motivation for a systematic design of the assembly system, has therefore become an important tasks to handle at FHP. The project continues, and the company has become aware of the importance of focusing on the system design as well as on the product development process. In the long run, this may give the preconditions for a better product and system development.

6. ACKNOWLEDGEMENT

Financial support for the development and the application of the method has been received from the National Swedish Board for Technical Development. Employees at FHP Elmotor AB have contributed to the research project by their involvement in the described assembly system design project.

REFERENCES

Bowen, H.K. and K.B. Clark (1994). The development projects: The Engine of Renewal, *Harvard Business Review*, Sept/Oct.

Feldmann, K. and T. Krebs (1990). Planning Programmable Assembly Systems, *11th Int. Conf. on Assembly Automation*, Michigan.

Lundström, M. (1993a). *Assembly Systems - An Approach to a System Design Method*, Thesis No.387, LiU-Tek-Lic-1993:29, Linköping.

Lundström, M. and C. Johansson (1993b). A Planning Method for the Design of Flexible Assembly Systems. *Conference proceeding of the 12th Int. Conf. on Production Research*, Lappeenranta, Finland.

Maczka, W.J., ed. (1985). Planning for Automated Assembly: The common Denominators, *Assembly Engineering* **2/85**.

Milberg, J. and M. Schmidt (1990). Flexible Assembly Systems-Opportunities and Challenge for Economic Production, *Annals of the CIRP*, Vol. **39/1**.

Oliva, L.M. (1990). Teamwork as a Competitive Advantage. *Conference proceedings of Autofact 90'*, Detroit, Michigan.

Robson, R.E., (1990). *The quality and Productivity Equation, American Corporate Strategies for the 1990s*, Productivity Press, Cambridge. Ed.

Skinner, W. (1978). *Manufacturing in the corporate strategy*, John Wiley & Sons, New York.

Warnecke, H.J., M. Schweizer and J. Schöningen (1987). *Taktzeitprognoze für flexible Montagestationen*, VDI-Z Bd. 129, **Nr. 3 März**.

Wu, B. (1992). *Manufacturing Systems Design and Analysis*. Chapman&Hall, Great Britain.

Yin, R.K. (1989). *Case Study Research, Design and methods*, Applied Social Research Methods Series, **Vol. 5**, London: Sage.

SUPPORT IN SETTING FEED RATES AND CUTTING SPEEDS FOR CNC MACHINE TOOLS THROUGH OVERRIDE LOGGING: PRACTICAL TEST RESULTS WITH A NEW CNC COMPONENT

Sören Striepe

Institut für Arbeitswissenschaft der Universität Gesamthochschule Kassel
Heinrich-Plett-Straße 40, D-34109 Kassel, Germany

Abstract: This study looks at technical ways of supporting the defining of machining rates on CNC machine tools. It is founded on an analysis of 128 skilled workers' work action in planning and optimizing machining rates when compiling and running-in NC programs. The study serves as a basis for the further development of CNC systems. Recommendations are made for the design of the human-machine interface of future CNC machine tools. These recommendations are then exemplified in a new CNC component called override logging which was designed and tested with skilled workers both in a laboratory and in a practical industrial setting.

Keywords: Manufacturing processes, CNC, Machining, Logging, Human-machine interface

1. DESCRIPTION OF THE PROBLEM

CNC technology is the most widespread computer technology in industrial manufacturing. CNC machine tools are the most frequently used and most developed CIM components (Martin and Rose (Eds.), 1992). Important areas of application for CNC systems are in milling, drilling and turning. 70% of all machine tool controls are implemented in these production processes (Krause, 1991). The most important movement parameters of the cutting tools are the feed and cutting rates which run at defined speeds and which are called machining rates.

Defining optimal machining rates for a given job situation is important with respect to the economic use of company resources. These rates directly influence the piece's quality, the production time and the wear on the production facilities.

Machining rates are set by means of a planning and optimization process while the program is being compiled and while running-in. This is done by a specialized programmer or by the skilled worker working the machine. These rates must be planned and programmed as part of the NC program before actual machining begins. Through the increasing popularity of shop floor programming (see Nuber and Schultz-Wild, 1990) the planning and input of machining rates are more and more often undertaken directly on the machines by the shop personnel. Furthermore, one can observe that when setting up, preplanned machining rates must often still be adapted and optimized according to the actual machining conditions.

Present available resources for defining machining rates, such as reference value charts, tool catalogues and indexes or computer aided machining rate data and knowledge-based machining data processors, concentrate on support before actual work on the floor. Few presently marketed computerized numerical controls (CNCs) offer options allowing the comfortable optimization of machining rates while running-in. Technical improvements with respect to defining machining rates while machining have been called for by all involved with the development and application of CNC machine tools and who wish to achieve optimal use of company resources (i.e. users, developers and builders of CNC systems; production and manufacturing scientists; ergonomists) (see i.e. Warnecke and Mertens, 1988; Hekeler, 1988; Bullinger and Ammon, 1991; Böhle, 1992; Martin and Rose (Eds.), 1992; Storr, et al., 1993; Schulz and Fechter, 1994). There is extensive consensus on the fact that technical improvements must be orientated by the work action of the machine operators (usually for CNC machine tools qualified skilled workers) in order to allow their experience based knowledge to be incorporated into the process of defining rates.

2. GOALS

This study looks at technical ways of supporting skilled workers in defining machining rates on CNC machine tools (Striepe 1995). Based on an analysis of the work action involved in planning and optimizing feed rates and cutting speeds when compiling and running-in NC programs, the study provides a basis for the further development of CNC systems. Recommendations are made for the development, design and realization of future CNC machine tools. These recommendations are then exemplified and tested in the form of a new technical option which logs manual feed rate and cutting speed adjustments by means of the override functions and which transfers these adjustments to the NC program.

3. PROCEDURE

3.1 Scope of the study

This study was carried out as a part of the CeA-Forschungsverbund (Research Association for Computer Supported Experience-Guided Work) and promoted by the German Ministry of Research and Technology (BMFT) and several builders of CNC systems (Martin (Ed.), 1995). This research association investigated in close interdisciplinary cooperation with ergonomic and engineering research institutions, work action on CNC machine tools when turning and milling and elaborated certain basic points for further technical development. All essential activities of the skilled worker while working on CNC machine tools (i.e. planning, programming, setting-up, running-in, optimizing and monitoring the process) were observed. The part of the study described in the following concentrates on specific aspects involving the defining of machining rates. The new override logging component was developed in close cooperation with the Institut für Arbeitswissenschaft der Universität Gesamthochschule Kassel (IfA-GhK), the Institut für Werkzeugmaschinen und Fertigungstechnik der Technischen Universität Berlin (IWF) and the Siemens AG Erlangen (Metzler and Striepe, 1993).

3.2 Analysis of work action

The prototypical development of an override log was preceded by a broad analysis of work action of skilled workers when defining machining rates on turning, milling and drilling CNC machines. A sample of 128 skilled workers from 30 different metal working industries were systematically observed and questioned. The sample was chosen as to reflect to a high degree the diverse areas of application of CNC technology in Germany with respect to:

- the structure of the companies (small-, medium- and large-scale enterprises)
- the spectrum of the machines and CNC systems used (51 machine types from 38 builders; 48 CNC types from 14 builders)
- the work organization (office programming, different forms of shop floor programming)
- the employee qualifications.

3.3 Practical Tests

The override logging tests were to be carried out in as practical conditions as possible and they were to lead to results that could be easily generalized. Thus, two prototypes were developed and tried in an early phase of development. These tests were carried out with skilled workers in order to be able to directly record the effects of the override logging on their work action, as well as to allow their expert knowledge to be incorporated at an early point. The skilled workers were observed in their work with the new component and then in light of the observations, questioned in partially guided interviews.

This graduated procedure made it possible to effectively try the override log. Closeness to actual practice was achieved by having the skilled workers carry out representative run-in tasks (in the lab - specifically defined; on the floor - day to day orders) on fully equipped CNC machine tools.

The practical tests were run by five very experienced skilled workers, who were experienced in using the technology involved (machine, CNC). First, a prototype was tested on a Boehringer VDF lathe with a Sinumerik 880T control and a separate PC (Fig. 1).

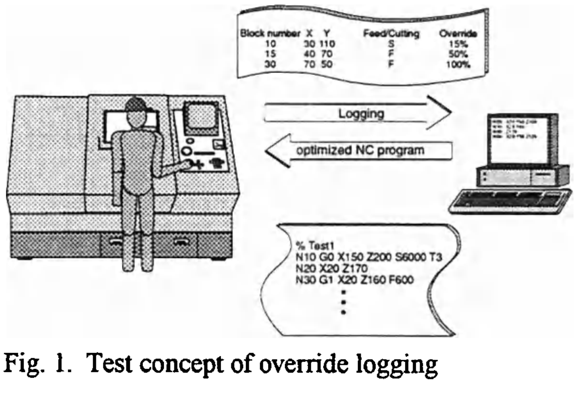


Fig. 1. Test concept of override logging

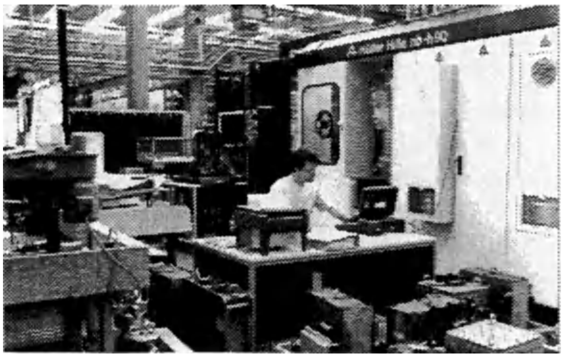


Fig. 2. Industrial test: Optimizing the NC program by means of override logging

With the help of the new component, each time the override switches are used, the override rates, the relevant block number and the actual tool position are transferred into a log on the PC. This data is then available as a basis for program optimization after running-in. For the first prototype, the dialog occurs using the Windows-principle and a mouse control. The optimized program can then be transferred back to the CNC for repeated runs.

The second prototype was installed on a Hüller-Hille machining center with a Sinumerik 880M CNC (see Fig. 2). This prototype functioned basically on the same principle as the other except for slight adaptions based on weaknesses identified in the first prototype. An important adaption was that the mouse controlled dialog was switched to the softkey dialog principle with function keys customary to CNCs.

4. A TECHNICAL CONCEPT FOR OVERRIDE LOGGING

The technical realization requires installing two new modules in the CNC (see Fig. 3).

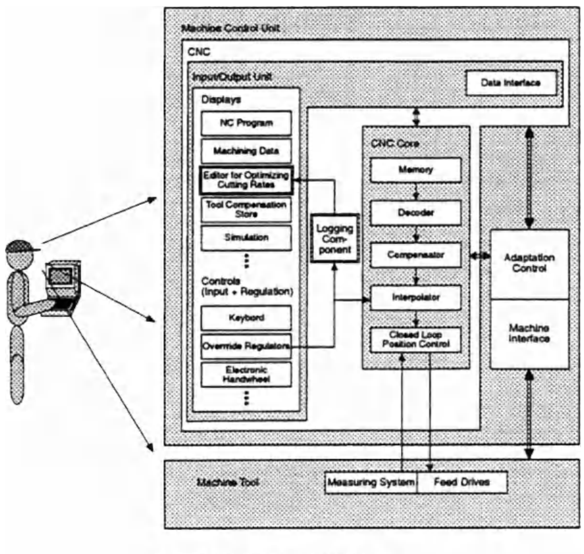


Fig. 3. New CNC modules for the realization of override logging

The first module is an override log book which logs the feed rates and cutting speeds, as well as the block number, G-functions and actual tool position. Each time the override switch is activated, the component logs these rates. It is essential to record the program position in order to be able to judge the regulated override rates with respect to the programmed course after running-in and in order to then change the rates in the NC program. Recording the actual tool position is important in order to mentally recreate the connection between manual changes and the machining process.

The second module is an editor which facilitates coordination between the data which is to be recorded and the NC program. The NC program and the log appear side-by-side on the screen allowing the workers to compare the data and judge the machining process and the rates in retrospect. Dialog functions for the adoption of regulated machining rates into the NC programs are provided (see Fig. 4).

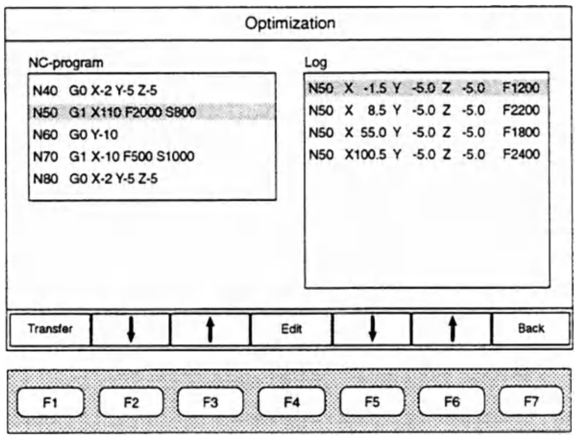


Fig. 4. Screen picture: Adoption of specific rates

5. RESULTS OF THE BASIC ANALYSIS OF WORK ACTION

The development of the override log was founded on the following points, which were the result of the broadly based analysis of work action in defining machining rates.

1. The machining rates inherent to the NC programs must normally be screened and optimized with respect to their adequacy for actual cutting conditions. Approx. 90% of the skilled workers check the programmed machining rates before or while running-in (see Table 1). 90% of the time, it is necessary to optimize the machining rates in the NC program (92% for at least every 10th program; 47% for at least every 2nd NC program) (see Table 2).

Table 1 Estimated necessity to check machining rates when running-in

	Number of specifications in %		
Estimated necessity to check machining rates	self-made NC programs (N = 78)	pre-made "foreign" NC programs (N = 60)	old, past-tested NC programs (N = 90)
always	91	98	86
mostly	6	2	6
often	3	0	4
seldom	0	0	2
never	0	0	2

Table 2 Frequency of adaptations to programmed machining rates

Estimation of the frequency of adaptations to programmed rates	Percentage of times mentioned (N = 69)
not at all	8
for 10 % of the programs	30
for 30 % of the programs	15
for 50 % of the programs	9
for 70 % of the programs	10
for every program	28

2. The override functions for the feed rates and cutting speeds play a central role when optimizing the cutting conditions. The rates are the machining conditions most often varied for process optimization (100% said they often optimize the process by changing the feed rate; 87% often change the cutting speed). In comparison, other means of optimization are less often employed (changes to the geometric parameters 45%; changes to the tools or machining order 15% respectively) (see Table 3). Most of the workers (82%) use the override functions instead of interrupting the run-in in order to change the rates in the NC program (see Table 4). The reason for this is that the override functions allow direct changes to be made to the machining process, making process changes immediately ascertainable.

Table 3 Often changed machining conditions

Machining conditions specified, which are often changed when running-in	Frequency of answers in % (N = 83)
feed rate	100
cutting speed	87
geometrical course, depth of cutting	45
cutting tool	14
order of machining steps	11
workpiece chucking/clamping	1

Table 4 Measures taken in optimizing rates

Specified measures taken in optimizing feed rates and cutting speeds when actually machining	Frequency of answers in % (N = 97)
changes to the rates using the override functions	82
changes to the rates in the NC program	18

3. Those machining rates optimized by using the override functions represent an important basis for further machining runs and should normally be transferred into the NC programs. In order to accomplish this, the operators must laboriously interrupt the run-in and then start up again once the changes to the program have been made, or they must note the relevant parameters to be changed (regulated machining rates; program positions; geometric coordinates).

Replacing already programmed machining rates with new rates is troublesome for the following reasons:
- It is not always possible to completely note the relevant parameters.
- Noting the parameters disturbs concentration on process optimization.
- Interruptions to the process disturb the run-in and cost time.
- Dialog operations and geometric definitions for program changes are very complicated, especially if they necessitate progressive graduations in the assignment of machining rates to geometric segments of the machining course (i.e. several feed rates for one machining stage in one cycle).

These procedures cost time and are open to mistakes. For these reasons, most of the skilled workers questioned (80%) expect support in their work from a technical option for logging manual override changes and for their adoption into the NC programs as described to them in the interviews (see Table 5).

Table 5 Expected support through override logging

Estimation of support of work action	Frequency of times mentioned in % (N = 79)
workers expect support	80
workers expect no support	20

6. RESULTS OF THE PRACTICAL TESTS

6.1 *Support effect of override logging in defining machining rates*

The practical tests showed that override logging would, as a component in future machine tool controls, have a very positive effect on defining machining rates. This holds for the following points:
- the reduction of disturbances while running-in
- support in recalling and judging the machining course and one's own override operations
- the reduction of the effort necessary when entering and calculating changes of machining rates
- support in planning machining rates
- support in accomplishing concrete order specifications (workpiece quality, production times, wear and tear on production facilities).

1. The reduction of disturbances while running-in results from:
   - a decrease in necessary run-in interruptions
   - a decrease in the effort necessary for noting (mentally or written) the machining parameters relevant when optimizing the programmed rates.
   - temporary postponement of possible changes to the original program until the "quiet" working phase after running-in (while running-in, concentration is necessary in order to avoid collisions and for manual optimizations of the cutting conditions).

2. Support in recalling and judging the machining sequence and in one's own use of the override when changing the program after running-in is attained by:
   - logging relevant NC block numbers, manually adapted rates and the actual tool positions
   - the depiction of these data in connection with the basic program.

3. A reduction in the effort involved in changing the machining rates inherent to the program results from a decrease in input operations and necessary coordinate calculations.

4. Support in the planning of machining rates when drawing up programs results from the reduced necessity to exactly preplan rates due to improved optimization possibilities while running-in.

5. The above described effects ultimately result in support in the realization of concrete order specifications due to:
   - a reduction of production times (especially 15-30% of the machining time and 10% of the run-in time)
   - a reduction and calculability of wear on the production facilities (especially tool wear)
   - fulfilment of quality demands by means of easier procedures and improved procedure options.

6.2 Design recommendations for future realizations

Several design recommendations for future realization, of this component as a mass-producible part of CNCs were arrived at by observing how the skilled workers used the override logging prototypes as well as in the interviews. These recommendations concern:
- the override log functions
- the design of screen pictures, information and dialogs
- the integration of the override log into machine tool controls.

These recommendations are briefly described in the following. For a more detailed description see Striepe (1995).

1. Side-by-side picture of the NC program and logging parameters in the override logging or control editor. This serves as a basis for the optimization of the machining rates in the NC program.
2. Scroll functions for NC program windows, logging windows and input fields as a means for orientation and for the selection of specific information (lines, parameters).
3. Logging of the information necessary for program optimization:
   - NC block number
   - actual tool position at the moment the override is used
   - regulated override values for feed rates and cutting speeds
   - the parameters of the override applications which were employed in the time between two blocks (during single block)
   - automatically produced demarcations between override regulation fields
   - personal notes made by the worker for logging sequences judged as being particularly important.
4. Technical options ensure comprehensibility and clarity of the logging graphics. These options:
   - highlight logging sequences judged important
   - illustrate the logging development during the run-in
   - specifically control the logging (reducing the log volume)
   - optimize the program periodically (reducing the log volume - see 9.) and
   - avoid redundant depiction of logging information.
5. Clarity of the parameters specific to the machining rates through an information window which can be called up as a whole and which supplements the NC program window.
6. A simple dialog for the substitution of programmed rates with singly logged or specifically selected rates. It is possible to directly enter one's own rates without using the input field.
7. The easy adoption of several logging rates for the graduated setting of machining rates in the NC program by means of:
   - system generated suggestions of NC blocks drawn from the program and logging parameters
   - comfortable ways of modifying the suggested NC blocks
   - options for the direct and complete adoption of all the generated NC blocks.
8. Direct feedback about program changes appears on the screen and access to the original program data by means of an intermediate data base.
9. An option for program changes using the override log in purposefully chosen run-in interruptions.
10. Compatibility of the override log and the CNC with respect to screen pictures, dialog technology and dialog design.

11. Support in the use of the override log through the integration of its modules into the CNC.
12. Reduced run-in interruptions while logging by means of much larger regulation ranges inherent in the override functions (at least 200%).

7. CONCLUSION AND ASSESSMENT OF THE RESULTS

The results of this study represent a basis for the further development of CNC machine tools. The description of machine operators' work action and of the deficits of present CNC machine tools, offers manufacturers a detailed picture of actual problems and necessary improvements. This study could serve as a basis for the development of specific recommendation manuals etc. It could also prompt further and even new technological developments.

The override logging results, backed by the prototypical realizations and the practical tests, are a very concrete basis for the development of the log as a component in marketable CNCs. System builders can draw directly on the potential support offered by the override log, as well as on the design recommendations for functions, ergonomic design and control integration.

Due to the small number of tests, the results of the practical tests of the override log cannot be backed statistically. However, the results link with the preceding broad study of work action when defining machining rates, which in itself is quite representive. The results of that study formed the empirically founded basis of the draft and prototypical design of the override log.

The results of the practical tests can basically be extended to all ranges of application of CNC machine tools for turning, drilling and milling. The tests showed no principle limits to the applicability of the override log in optimizing feed rates and cutting speeds. Although the log was tested on a Sinumerik 880 control, it can, in principle, be technically realized in the same way on other types of CNCs.

That the results of this study are practically oriented and relevant is emphasized by the fact that the largest European builder of machine tool controls - the Siemens AG - is presently working on making the override log part of their control package, based on the analysis and recommendations presented here. The advantageous effect of the override log on both the work action and economic production, as shown in the practical tests, motivated this CNC builder's decision. The resulting design recommendations act as a guide for further development. The fact that customers are already demanding the override logging component announced by the Siemens AG's sales department, shows that the technical support developed in this study anticipates practical problems and thus market demands.

REFERENCES

Böhle, F. (1992). Einfahren und Optimieren. In: *Erfahrungsgeleitete Arbeit mit CNC-Werkzeugmaschinen und deren technische Unterstützung* (Institut für Arbeitswissenschaft der Universität Gesamthochschule Kassel, Ed.), pp. 46-51. Institut für Arbeitswissenschaft, Kassel.

Bullinger, H.-J. and R. Ammon (1991). Werkstattorientierte Produktionsunterstützung. In: *Fertigungstechnisches Kolloquium am 1./2. Oktober 1991 in Stuttgart*, pp. 68-73. Springer, Berlin.

Hekeler, M. (1988). Werkstattorientierte Programmiertechnik für Facharbeiter. *Technische Rundschau* **80 (1988) 38**, pp. 52-60.

Krause, N. (1991). Steuerungssysteme für Werkzeugmaschinen und Industrieroboter. In: *Fertigungstechnisches Kolloquium am 1./2. Oktober 1991 in Stuttgart*, pp. 41-47. Springer, Berlin.

Martin, H., Ed. (1995). *CeA: Computergestützte erfahrungsgeleitete Arbeit. Integrierte Arbeitsgestaltung und Technikentwicklung als Innovationsstrategie*. Springer, Berlin.

Martin, H. and, H. Rose, Eds. (1992). *CNC-Entwicklung und -Anwendung auf Basis erfahrungsgeleiteter Arbeit*. Wirtschaftsverlag NW, Bremen.

Metzler, U. and S. Striepe (1993). Override-Protokollierung - eine neue CNC-Funktion. *Zeitschrift für wirtschaftliche Fertigung und Automatisierung ZwF* **88 (1993) 9**, pp. 392-394.

Nuber, Ch. and R. Schultz-Wild (1990). Facharbeitereinsatz und Verbreitung von Werkstattprogrammierung. In: *Programmieren in der Werkstatt* (H. Rose, Ed.), pp. 155-183. Campus, Frankfurt/Main.

Schulz, H. and Th. Fechter (1994). Defizite der heutigen Werkstück-Programmierung. *Werkstatt und Betrieb* **127 (1994) 1-2**, pp. 18-21.

Storr, A., W. Kempf and C. Itterheim (1993). Werkstattgerechte Nutzerunterstützung für die Freiformflächenbearbeitung. In: *Werkstattgerechte Nutzerunterstützung für die Freiformflächenbearbeitung* (D. Kochan, Ed.), pp. 20-34. VDI, Düsseldorf.

Striepe, S. (1995). Technologiewertbestimmung an CNC-Werkzeugmaschinen: Anforderungen an die technische Unterstützung bei Programmerstellung und Einfahren. Overrideprotollierung. Dissertation im Fachbereich Maschinenbau der Universität Gesamthochschule Kassel.

Warnecke, G. and P. Mertens (1988). Aufbau und Anwendung von wissensbasierten Technologieplanungssystemen. *Zeitschrift für wirtschaftliche Fertigung und Automatisierung ZwF* **83 (1988) 11**, pp. 559-562.

TEACHING MOTION/FORCE SKILLS TO ROBOTS BY HUMAN DEMONSTRATION

S. LIU and H. ASADA

Massachusetts Institute of Technology, Center for Information-Driven Mechanical Systems, Department of Mechanical Engineering, Cambridge, Massachusetts, USA

Abstract. In this paper we survey a number of methods for teaching human motion/force skills to robots. Since verbal description of a human skill could only provide qualitative guidelines for machine programming, several works were motivated to identify human skills analytically through the study of human teaching data. The key advantage of direct teaching by human demonstration is that the method is highly user-friendly; it requires no special knowledge of machine programming from the users, and it automatically extracts user intentions and strategies from task performance demonstrated by the users. From the early teaching/playback method to the recent methods of teaching sensory feedback laws, a historical perspective of the field of human skill teaching is given in this paper. Limitations and applicability of methods under review are discussed. Essential issues such as representation and identification of human skills as well as consistency of teaching information are addressed.

Key Words. Robot programming; robot teaching and learning; human skill transfer; feedback control; adaptive control

1. INTRODUCTION

Skill is the ability to perform a learned physical task *and* the intelligence to cope with uncertainties in the task based on sensory perception. Humans are uniquely capable of manipulating objects with skills to accomplish complex tasks. However, most people cannot describe precisely how they manipulate objects or interact with the environment. The detail of this ability and intelligence is invariably lost in the human subconscious, thereby making it highly difficult to quantify human skills and code them into machine instructions.

There have been a number of attempts toward developing human-like skills for robotic systems based on various approaches, such as task-level machine learning, teaching-by-showing, and the recently developed perception-action models of human skills. Task-level machine learning, such as the one proposed by Aboaf *et al.* (1989), requires the use of an accurate task model and an iterative learning process to build up machine intelligence. The result of learning is task specific, and its efficacy depends heavily on the model accuracy. In this paper, we focus on teaching methods for transferring human intentions, strategies, and skills to robots based on human demonstration. The key advantage of direct teaching of human skills is its user-friendliness in robot programming: it requires no special machine programming knowledge and allows factory floor workers to easily comprehend and effectively use sophisticated robots. In view of today's popular just-in-time strategy for manufacturing, this is particularly desirable for meeting small-batch, frequently changing task demands. In addition to this practical aspect, study of human skill is also essential to developing the theory of machine intelligence.

This paper gives a comprehensive discussion on the evolution as well as the state of the art in teaching motion/force skills to robots, and is organized as follows. In Section 2, a historical perspective on this area of study is given. In Section 3, essential issues in human skill acquisition, in particular, representation and quantification of tool/object manipulation skills are addressed. Research opportunities in teaching of human skills as well as their relevance to other fields of study are discussed in Section 4. Concluding remarks are given in Section 5.

2. HISTORICAL PERSPECTIVE

The advancement in teaching human motion/force skills to robots has been parallel to the development of general approaches to robot control and programming. This is a natural result since only when the framework of robot control and programming was better established, did one realize more about what to look for in human demonstration that could be useful to robots.

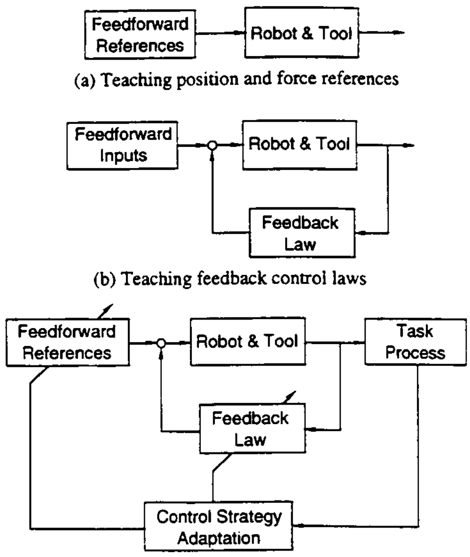


Fig. 1. Progress in teaching human skills to robots.

Lozano-Pérez (1983) considered that general approaches to robot programming fell into three broad categories: *teaching-by-showing*, *robot-level programming*, and *task-level programming*. Here we review teaching methods for human skill transfer from a control engineering viewpoint, with the intention that these methods can be extended to a broader class of control system problems. The progress of teaching methods for transfer of human motion/force skills in history can be roughly divided into the following stages: 1) teaching reference position trajectories, 2) teaching reference contact forces, 3) teaching feedback control laws, and 4) teaching task-level adaptive control laws. Using block diagrams typical for general control systems, these stages are depicted in Figure 1. The historical background and the representative works in these stages are discussed in this section.

Teaching Reference Trajectories

In 1954, George C. Devol filed a U.S. patent for a new machine for part transfer, and he claimed the basic concept of *teaching/playback* to control the device. In principle, this teaching/playback method is a way to convey human intention in terms of position trajectory to robots by manually moving the robot arm through the desired points in work space. This method is highly user-friendly, and therefore has its widespread applications in the early industrial robots, such as pick-and-place, spray painting, and spot welding. However, by teaching/playback, only a single execution sequence of robot motion is specified; no sensory information has been incorporated into any feedback loop to account for uncertainties in the task environment. Clearly this deficiency limits this method to only simple tasks where uncertainty is minimal and no sensory feedback is needed.

Teaching Reference Forces

Robot force control began with telemanipulator and artificial arm control in the 1950s and 1960s (Whitney, 1987). In these applications, humans are directly involved in the control loops, through master-slave teleoperation or direct connections between joint motors and muscle electrodes. Therefore, a desired contact force with the task environment can be directly specified by the human operator.

For autonomous robots, Asada (1979) presented the force teachable robot such that the robot can interact with external objects or environment with a prescribed contact force. In that implementation, the robot consisted of an arm and a hand with elastic fingers. The arm's movement was controlled with the position teaching/playback scheme, while the hand grasped objets with forces also taught by a human operator. The teaching of reference force is analogous to the conventional teaching of reference position, therefore it is also highly time-effective for robot programming. Hirzinger and Heindl (1983) designed a force-torque-sensor ball which can be used to teach a robot reference paths for free motion in space and reference forces/torques when in contact with environment.

These two works represent the early research on direct teaching of position and force references to robots. Around the same time compliance and force controls in robotic systems were active research topics in the literature (Hanafusa and Asada, 1977; Hogan, 1985; Mason, 1981; Salisbury, 1980). The importance of specifying desired force-motion relations to robots interacting with task environment was clearly identified from these works. Hence, teaching of reference force-motion relations became the focus of study thereafter.

Teaching Feedback Control Laws

One of the main challenges in robot programming arises from uncertainties in task environment with which robots interact. Using sensory feedback in control, robots can cope with a greater degree of uncertainty than without any sensing. The central issue in robot programming is therefore the design of a mapping (function) in the computer program that generates effective corrections in robot motions given a certain sensory feedback information (Lozano-Pérez, 1983). Obviously this is related to design of feedback control laws for more general control problems.

In the context of teaching-by-showing, related research has been focused on extracting or recovering a mapping, from human demonstration data, that associates human responses with stimuli. This mapping, in an explicit mathematical or symbolic form, can then be used as an equivalent to a feedback control law in robot control systems. In addition to position information, common sensory signals in robotic applications also include force, vision, and tactile. For the case of force feedback, the feedback control law is taught so that a robot's closed-loop behavior approximates a desired *compliance* or *impedance*, which defines an explicit relation between reaction force and robot motion. Hirzinger and Landzettel (1985) showed that through on-line corrections induced by a human teacher via a force-torque-sensor-ball, a robot can be taught with a desired stiffness in its interaction with the environment. An associative memory was introduced to store a set of corresponding input (sensory signal) and output (robot motion) data such that during playback similar inputs mapped to similar outputs. The set of data used for training the associative memory was obtained during robot playback from human supervisor's corrections. The human's intention and strategy was conveyed to the robot only indirectly through a sensor ball. Since the human supervisor had not given any demonstration of his or her skill under a natural situation, the essential skill, such as the tool holding impedance in a deburring task, could not be extracted through this indirect teaching manner.

Asada and Izumi (1989) presented a method for automatic program generation for the hybrid control of robots based on human teaching data. In principle, this is a direct approach to teaching a human operator's skill in terms of force-motion relation in a task with force and motion constraints, since a human actually demonstrated the task performing skill in a natural manner during teaching. Asada and Asari (1988) applied this direct method to the case of deburring where the tool holding impedance of an experienced worker was identified from actual demonstrations.

For the case of using vision as the feedback signal, Kuniyoshi *et al.* (1990) proposed a theory for visual recognition of human action sequences, which can be used for direct transfer of implicit task knowledge from human operators to robots. In this work, vision was used for both tracking movement of a human hand for program generation and for on-line feedback during task execution by the robot. Ikeuchi and Suehiro (1992) devised a similar approach to automatic program generation for robot assembly task, and established the paradigm termed *assembly plan from observation.* Takahashi *et al.* (1993) also presented a similar method for analyzing human assembly operations for use in automatically generating robot commands.

Yang *et al.* (1993) proposed a hidden Markov model (HMM) approach to human skill modeling and transferring. The HMM method treated its observation on a symbolic level, and that made the fusion of different sensory signals possible regardless of their physical meanings. With the framework of HMM, a *most likely criterion* is defined as a quantitative measure based on which the underlying skill can be uncovered from a set of human performance measurements. The concept of most likely performance makes it possible to use stochastic methods to cope with uncertainties in both human performance and environment. This has been demonstrated in (Yang *et al.*, 1993) for teaching reference position trajectories. Application of the HMM method to transferring feedback control laws is currently under investigation.

So far, the *teaching-by-showing* method has been extended from teaching a position or force reference to a feedback control law that defines a mapping between sensory feedback inputs and robot motion outputs. This mapping, in most cases an associative memory, can be constructed based on human teaching data taken either from human-supervised robot motion or natural human motion (demonstrations). Although this mapping was readily applicable to robot control systems for generating human-like skillful motion, it was later realized that this mapping defined skill only under a certain restricted task condition. When the task condition varied substantially, the mapping relationship found in a human's motion also changed accordingly. This problem was formulated by Liu and Asada with the framework of adaptive control (Liu and Asada, 1992; Liu and Asada, 1993), and teaching-by-showing was further generalized to transferring task-level adaptive control laws from humans to robots.

Teaching Task-Level Adaptive Laws

Liu and Asada (1992) considered skill as the generic ability to perform a class of tasks; that included association and generation of action strategies based on the perception of the task process state. They interpreted this perception-action cycle as a kind of adaptation capability in performing task with uncertainties and variations. Hogan (1992) observed a similar hierarchical structure in human movement production with a generic plan formulated at an abstract, task level and implemented at a more concrete continuous level. Using such models for human manipulation motion, the target skill to be taught to robots should include not only the control strategies for machine-level, continuous movement, but also the task-level adaptation law or a generic strategy planner. The hierarchical structure involving an inner feedback loop and an outer adaptation or strategy planner loop is depicted in Figure 1.

In representing the task-level adaptive law, Liu and Asada (1993) used a neural network to store the associative memory that related task-level process parameters to control strategy parameters. More specifically, using deburring as the example task in (Liu and Asada, 1993), the task-level process parameters referred to burr size and hardness and were identified based on a deburring process model, while the control strategy parameters included tool feedrate and tool holding compliance. In (Shimokura and Liu, 1994), a laser gap sensor was utilized for direct burr size measurement in identifying the associative memory from a human. Yang and Asada (1992) employed a similar control structure in which a set of linguistic control rules was used to represent the associative memory obtained from human description. McCarragher (1994) applied qualitative reasoning to interpret force signals generated by humans, and a discrete event controller for task-level decision making in a assembly process.

To recover adaptive control laws from human teaching data based on the perception-action model, it is invariably necessary to identify the essential features in a task process that a human perceives when generating action strategies. If these features or process characteristics were not properly defined and parameterized, it is almost hopeless to recover a consistent mapping that represents the perception-action cycle from human teaching data. As pointed out in (Liu and Asada, 1992), consistency is a critical issue in human skill study, particularly with the attempt to extract useful information from human motion for robot programming.

3. ELUCIDATION OF HUMAN SKILLS

The common approach to human skill transfer for robot teaching has been generally focused on identifying a mapping (associative memory) that relates a human's perception of task process state to some effective control actions. From the viewpoint of control engineering, this mapping may manifest as a certain feedback control law in a human's motion, or as a high-level adaptive law that modifies feedback control to meet complex task environment. Within this framework, consistency arises as an issue which is critical to guarantee that every point in the perception space maps to at most one point in the control action space through the identified mapping. This is a necessary property for a skill since an inconsistent mapping would

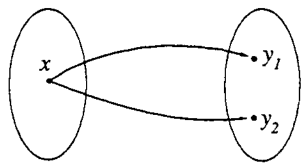


Fig. 2. Example of inconsistent relationship.

lead to undecidable cases where a given process condition dictates two or more different (possibly conflicting) control actions. In such cases, it is impossible to decide which action to take based solely on data points in the input space. Figure 2 shows an example of inconsistent relationship. A point defined by $\boldsymbol{x}$ in the input space corresponds to two distinct points in the output space, $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$. Note that a valid mathematical mapping can not be defined for this inconsistent relationship.

Since a human with a skill always make a clear control decision based on some perception inputs, it can be argued that the mapping from his/her perception space to his/her control action space is consistent. In fact, this mapping that mathematically characterizes a human skill is also a continuous function of some degree of smoothness. For the associative memory of a human, a small deviation $|\Delta\boldsymbol{x}| < \delta$ in the perception variables will only lead to a small deviation $|\Delta f(\boldsymbol{x})| < \epsilon$ in the desired control command. Namely, a human slightly modifies an action in response to a small change in his/her perception (Kosko, 1992). Therefore, uniform continuity over the entire input space is often a desired property in modeling human behavior.

Given a set of human teaching data, the first step toward skill transfer is therefore to construct a perception space and a control action space such that the perception-action map is consistent and continuous. To this end, a quantitative method is needed to evaluate the consistency and continuity properties even *before* the mapping function is identified. Liu and Asada (1992) devised a quantitative measure based on Lipschitz's condition which is widely used in calculus for evaluating and defining continuity of analytical functions. Lipschitz's condition states that a function $f(\boldsymbol{x})$ is uniformly continuous over its domain $X \subset R^n$ if:

$$\frac{|f(\boldsymbol{x}_2) - f(\boldsymbol{x}_1)|}{|\boldsymbol{x}_2 - \boldsymbol{x}_1|} \leq L \quad \forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in X \qquad (1)$$

Lipschitz's condition measures the ratio of the distance between two points in the output space over the distance in the input space: the function fails the test unless the ratio of those distances is bounded for all points in the domain of the function.

A similar test can be applied to sets of training data obtained from human demonstration. Let us consider a relation defined by a set of input-output pairs

$$\mathcal{R} = \{(\boldsymbol{x}_i, \boldsymbol{y}_i); \boldsymbol{x}_i \in X, \boldsymbol{x}_i \in Y\} \quad (2)$$

and the relation is assumed to be in a function form:

$$\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x}) \quad (3)$$

We can apply a Lipschitz-like condition by taking the ratio of the distance between two points in the output space over the distance separating them in the input space. If the ratio between any two points in the set $\mathcal{R}$ is above some arbitrary threshold value, then the condition is not satisfied. This Lipschitz-like condition is a very useful tool in the context of skill transfer since it can determine if a relation defined by a representative set of points contains inconsistencies or discontinuities.

An experienced human always performs a task successfully and without hesitation. This suggests that the mapping from the human's perception space to its control command space must also satisfy Lipschitz's condition. If some pair of data points in the training set $\mathcal{R}$ violates Lipschitz's condition, this implies the input space X for the training set $\mathcal{R}$ does not completely characterize human perception of the process state. Namely, there are some critical process features missing from the vector of input variables $\boldsymbol{x}$. This can easily happen when the input variable vector $\boldsymbol{x}$ includes only direct sensors signals, e.g. force and position signals, while a human relies on some process information other than just force and position to make a control decision. This situation can be interpreted geometrically as shown in Figure 3. The missing critical information lies along a dimension orthogonal to the original input space. Two points that are hardly separable in the input space (Figure 3-(a)) may be far apart along the missing dimension (Figure 3-(b)), which is how a human differentiates the two input points and thereby takes two distinct control actions. By introducing new dimensions to the input space, a set of properly selected additional variables could make the training data pairs satisfy Lipschitz's condition. In (Dupuis *et al.*, 1992; Liu and Asada, 1993), Lipschitz's condition was applied successfully to assembly and deburring tasks where consistent mappings were identified from human teaching data based on proper definition of perception spaces.

The Lipschitz test proved to be effective in examining the completeness of a given perception space. However, the use of the Lipschitz test is based on the assumption that human performance

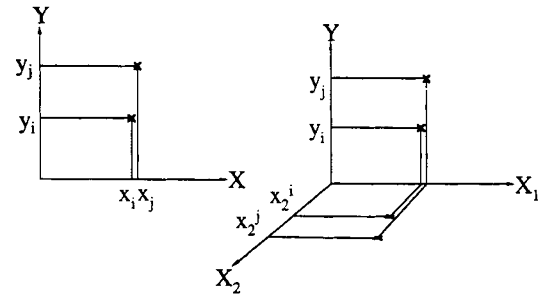


Fig. 3. Geometric interpretation of inconsistencies due to missing input information.

is consistent in terms of a certain target skill to be recovered from demonstration data. Physical data obtained from human demonstration inevitably could contain inconsistent information due to either measurement noise or human inconsistency. In (Liu and Asada, 1993), these inconsistent information was removed efficiently by using the Lipschitz test. In (Delson and West, 1994), Delson and West proposed an alternative method for interpreting human demonstration data. There the presence of human inconsistency in demonstration data was used advantageously to provide additional information regarding task requirements as well as strategies. One of the main features in Delson and West's work is that multiple demonstrations of the same task by the same person can be combined so that a more robust task strategy can be identified. This idea was demonstrated in (Delson and West, 1994) for a obstacle avoidance manipulation.

4. PROSPECT OF FUTURE RESEARCH

As discussed in the previous section, the identification of perception space is a critically important issue in human skill transfer. To teach human manipulative skills, it is necessary to know which pieces of information humans use and which particular aspect of the process they pay attention to. To represent a human skill, we need to identify a complete set of variables describing the information that the human perceives during the process. Based on the complete description of the perception space, the human skills and strategies can be represented properly and consistently.

This problem is a type of system structure identification, which is difficult to deal with. If the perception-action map describing a human skill is a linear map, the problem is rather simple; standard techniques based on identification error can be used to reconstruct the input space. Human skills, however, are highly nonlinear, hence traditional methods do not apply. The Lipschitz test described in this paper is a simple way of examining the completeness of a given perception space, and is applicable to nonlinear maps. The Lipschitz test, however, cannot be applied to noisy data or the ones containing inconsistent data. There is a way of alleviating the noise problem (He and Asada, 1993), but it is still an open question. Further investigation and the development of more powerful methods are needed to deal with inconsistent human data.

Studies on human skill teaching/transfer are not only useful for practical applications but also significant and essential for robotics research. Transferring human skills entails basic modeling and representation techniques, which are important issues in understanding manipulative strategies in general. Explicit representation and modeling of human behavior provide insights and manifest subconscious human skills and strategies. Moreover, direct observation and identification of human behavior, effective skills and novel strategies that have never been used in robotics can de discovered. Studies on human skill transfer may provide a new approach or a methodology for the study of manipulation and task strategies.

The issues central to human skill transfer are fundamental issues common to many other areas in robotics. In telerobotics and virtual reality, the communication of human intention and strategies as well as skills is a central issue. The broad field of human-machine systems and human motor psychology is directly related to human skill transfer as well. The utility of human skill transfer involves both practical applications and scientific importance, having significant connections to many related areas in robotics and beyond.

5. CONCLUSIONS

The key advantage of direct teaching by human demonstration is that the method is highly user-friendly; it requires no special knowledge of machine programming from the users, and it automatically extracts user intentions and strategies from task performance demonstrated by the users. From the teaching/playback approach for the early industrial robots, direct teaching has evolved into a class of highly sophisticated methods for transferring not just position intentions from humans, but also sensory-feedback, motion control laws as well as task-level strategy planning and adaptation laws. Several works have demonstrated the efficacy of direct teaching of human motion/force skills to robots. However, these implementations are still somewhat task-specific. In particular, in the attempt to recover a perception-action map from human teaching data, there are not yet systematic guidelines for constructing an adequate perception space that thoroughly captures essential task-level features that a human observes in his/her perception-action mapping process. One previous work using the Lipschitz test is good only for analysis of consistency for a given perception space. Further quantitative methods are needed to examine human teaching data, not only for constructing a meaningful perception space, but also for identifying inconsistency in human behavior and other possible noises inherent in teaching data. Previous results also revealed that inconsistent data obtained from human demonstration could in fact be used advantageously to extract additional task information for more robust and accurate robot motion. This is also a direction worth further pursuit.

6. REFERENCES

Aboaf, E. W., S. M. Drucker and C. G. Atkeson (1989). Task-level robot learning: juggling a tennis ball more accurately. In: *Proc. 1989 IEEE Int. Conf. Robotics and Automation.* pp. 1290–1295.

Asada, H. (1979). Studies on prehension and handling of objects by robotic hands with elastic fingers. PhD thesis. Kyoto University.

Asada, H. and H. Izumi (1989). Automatic program generation from teaching data for the hybrid control of robots. *IEEE Trans. Robotics and Automation* **5**(2), 163–173.

Asada, H. and Y. Asari (1988). The direct teaching of tool manipulation skills via the impedance identification of human motions. In: *Proc. 1988 IEEE Int. Conf. Robotics and Automation.* pp. 1269–1274.

Delson, N. and H. West (1994). Robot programming by human demonstration: the use of human variation in identifying obstacle free trajectories. In: *Proc. 1994 IEEE Int. Conf. Robotics and Automation.* pp. 64–571.

Dupuis, E., S. Liu and H. Asada (1992). On the consistency of compliant control laws acquired from humans. In: *ASME Winter Annual Meeting.*

Hanafusa, H. and H. Asada (1977). A robot hand with elastic fingers and its application to assembly process. In: *1977 IFAC Sym. Information and Control Problems in Manufacturing Technology.*

He, X. and H. Asada (1993). A new method for identifying orders of input-output models for nonlinear dynamical systems. In: *Proc. 1993 American Control Conference.*

Hirzinger, G. and J. Heindl (1983). Sensor programming - a new way for teaching a robot paths and forces/torques simultaneously. In: *Proc. 3rd Int. Conf. on Robot Vision and Sensory Controls.* pp. 549–558.

Hirzinger, G. and K. Landzettel (1985). Sensory feedback structures for robots with supervised learning. In: *Proc. 1985 IEEE Int. Conf. Robotics and Automation.* pp. 627–635.

Hogan, N. (1985). Impedance control: an approach to manipulation, part i-iii. *ASME J. Dynamic Systems, Measurement, and Control* **107-1**, 1–23.

Hogan, N. (1992). How humans adapt to kinematic constraints. In: *Proc. of the 7th Yale Workshop on Adaptive and Learning Systems*. pp. 182–185.

Ikeuchi, K. and T. Suehiro (1992). Toward an assembly plan from observation. In: *Proc. 1992 IEEE Int. Conf. Robotics and Automation*. pp. 2171–2177.

Kosko, B. (1992). *Neural Networks and Fuzzy Systems*. Prentice Hall. Englewood Cliff, New Jersey.

Kuniyoshi, Y., H. Inoue and M. Inaba (1990). Design and implementation of a system that generates assembly programs from visual recognition of human action sequences. In: *Proc. 1990 IEEE International Workshop on Intelligent Robots and Systems*. pp. 567–575.

Liu, S. and H. Asada (1992). Transferring manipulative skills to robots: representation and acquisition of tool manipulative skills using a process dynamics model. *ASME Journal Dynamic Systems, Measurement, and Control* **114**, 220–228.

Liu, S. and H. Asada (1993). Transfer of human skills to robots: learning from human demonstrations for building an adaptive control system. In: *Proc. 1993 American Control Conference*. pp. 1414–1417.

Lozano-Pérez, T. (1983). Robot programming. *Proceedings of the IEEE* **71**(7), 821–841.

Mason, M. T. (1981). Compliance and force control for computer controlled manipulators. *IEEE Trans. Systems, Man, and Cybernetics* **SMC-11**, 418–432.

McCarragher, B. J. (1994). Force sensing from human demonstration using a hybrid dynamical model and qualitative reasoning. In: *Porc. 1994 IEEE Int. Conf. Robotics and Automation*. pp. 557–563.

Salisbury, J. K. (1980). Active stiffness control of a manipulator in cartesian coordinates. In: *Proc. 1980 IEEE Conf. Decision and Control*.

Shimokura, K.-I. and S. Liu (1994). Programming deburring robots based on human demonstration with direct burr size measurement. In: *Proc. 1994 IEEE Int. Conf. Robotics and Automation*. pp. 572–577.

Takahashi, T., H. Ogata and S S. Muto (1993). A method for analyzing human assembly operations for use in automatically generating robot commands. In: *Proc. 1993 IEEE Conf. on Robotics and Automation*. pp. 695–700.

Whitney, D. E. (1987). Historical perspective and state of the art in robot force control. *The International Journal of Robotics Research* **9**(1), 3–14.

Yang, B.-H. and H. Asada (1992). Hybrid linguistic/numeric control of deburring robots based on human skills. In: *Proc. 1992 IEEE Int. Conf. Robotics and Automation*.

Yang, J., Y. Xu and C. S. Chen (1993). Hidden markov model approach to skill learning and its application in teleroboticn. In: *Proc. 1993 IEEE Int. Conf. Robotics and Automation*. pp. 396–402.

DEVELOPMENT OF MACHINE-MAINTENANCE TRAINING SYSTEM IN VIRTUAL ENVIRONMENT

Tetsuo Tezuka, Ken-ichiro Kashiwa, Takuya Mitani and Hidekazu Yoshikawa

Kyoto University, Institute of Atomic Energy, Gokasho, Uji, Kyoto 611, Japan

Abstract: The periodical inspection of nuclear power plants is indispensable in their operations. However, it requires a lot of workforces with a high degree of technical skill in assembling and disassembling various sorts of machines in hazardous environment. In this study the authors are developing a machine-maintenance training system in virtual environment which can be used easily by beginners. In this paper the system configuration is outlined. And proposed are recognition algorithms of natural gestures, which differ with trainees, and also a method for representing assembly and disassembly procedures by using Petri net.

Keywords: Virtual reality, nuclear power plant, machine-maintenance, training system, dataglove, gesture recognition, Petri net.

1. INTRODUCTION

The periodical inspection of nuclear power plants is indispensable in their operations which need absolute reliability about their safety. However, it requires a lot of workforces with a high degree of technical skill in assembling and disassembling various sorts of machines in hazardous environment such as in a nuclear reactor.

In such machine-maintenance training of the workers the following points are important.

- The trainee feels being at the real plant.
- Models which look exactly like the machines on site are used in the training.
- Appropriate advices are given to the trainee by the supervisors when the trainee have made some mistakes.
- The assembly and/or disassembly procedure is shown whenever the trainee wants to see.

In this study the authors are developing a machine-maintenance training system in virtual environment.

The expected effects are :

- Any machines on site can be built in virtual space.
- The system can reduce the work-load of the supervisors in the training.
- The trainee can get the skills for maintenance efficiently compared with the training based on manuals and/or limited models.
- The system can offer the workers more chances to get sufficient training for the maintenance of various sorts of machines.

The virtual reality technology has recently been attracting much attention in variety of research fields. But the benefits of using virtual reality environment have not been fully investigated in spite of enormous cost for preparing computer systems, e.g. high performance computers, various sensors, a display for 3-dimensional images.

This study especially aims to investigate the problems in the following:

1. **How easily can the beginners use the training system without much instruction about it in advance?**

As is well known one can use only limited gestures of hands by using the dataglove system due to lack of the flex-sensors' ability. Therefore, some fixed gestures (or postures) are often used to send commands to computers in many systems developed in virtual environment. These commands are used to make the system development much easier, but are needless for the operations in the real world.

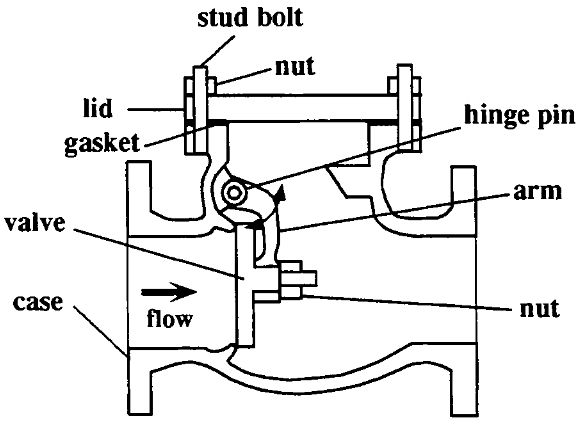


Fig. 1. Check-valve in section.

For example, there are many (dynamic) gestures for picking a nut. Having a special gesture correspond to a command, 'pick a nut', forces the operator to learn in advance how to use the training system. The system will be very difficult for the beginners to use without much instruction in advance, if the gestures recognized in the system are very limited (Papper, et al. 1993). Therefore, the system is required to recognize the natural hand-movements which a human is usually using in the real world and are different with persons. The system developed here includes the gesture recognition algorithms which recognize the meanings of hand-movements by using the information about the environment where the hand is doing something and also the action of the other hand.

2. **Is the virtual environment, e.g. 3-dimensional image and sound, useful for trainees to get skills required in assembling and disassembling machine?**

The effects of using virtual environment can be quantitatively evaluated by comparing the time to finish a given task in the real environment after training through various sorts of training systems. The qualitative evaluation such as comments given by trainees is also important.

In the next section, the basic tasks which should be represented in the system for machine-maintenance training are summarized by taking a check-valve as an example. In section 3, the system configuration is outlined, and some restrictions of the system due to the hardware characteristics are described. In section 4, the objects which are used for the system development are shown. And then, the recognition algorithms and the representation method of the various tasks by using Petri net are proposed in section 5 and 6.

2. BASIC TASKS IN MACHINE MAINTENANCE

2.1 *Disassembly procedure of check valve*

The procedure of disassembling a check valve shown in Fig.1 is as follows:

1. Mark both the lid and and the case to put the lid at the right position after disassembly.
2. Loosen all the nuts on the lid by a wrench in a right manner. The nuts should not be got rid of the bolts at this stage to avoid an accident, because the internal pressure of the check valve might remain very high.
3. Lift up the lid to confirm that the internal pressure be normal.
4. Get rid of all the nuts by a hand, and put them in an appropriate place. The latter task is omitted in the following.
5. Get rid of the lid.
6. Get rid of the stud bolts by double-nut method (See section 6).
7. Get rid of the gasket by a spatula.
8. Hold the arm or the valve and pull the hinge pin out of the arm.
9. Get rid of the valve and arm.
10. Separate the valve from the arm.

The assembly procedure is just in the reverse order.

2.2 *Basic tasks in assembly and disassembly procedures*

The basic tasks, which are necessary to represent the assembly and disassembly procedure in the virtual environment, are classified as follows:

1. Grasp (pinch) an object by a hand (or a tool).
2. Release a hand (or a tool) from an object.
3. Keep an object at a certain position.
4. Attach an object to another object.
   To put a nut on a table is represented by this task.
5. Separate the attached objects.
6. Put an object into a hole of another object.
7. Transfer an object grasped by something to a certain direction.
8. Rotate an object grasped by something around a certain axis.
9. Examine something about objects.

Let's take a task, 'loosen a nut and remove it from a bolt', as an example. This task can be decomposed into basic tasks as follows:

1. Grasp wrenches.
2. Attach wrenches to the nut and the bolt.
3. Turn the nut around its center by using the wrench.
4. Release the wrenches.
5. Pinch the nut and the bolt.
6. Turn the nut.
7. Get rid of the nut from the bolt.

This paper describes how to realize these tasks in virtual environment, that is, 3-dimensional space constructed by a computer, and also how to compose various complicated procedure of assembling and/or disassembling machines.

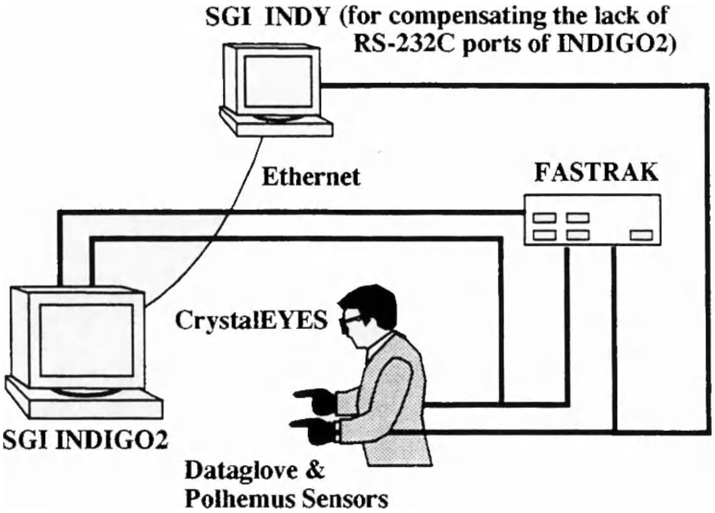


Fig. 2. Experimental system of machine maintenance.

3. SYSTEM CONFIGURATION

The hardware configuration of the training system developed in this study is shown in Fig.2.

In this system, the datagloves on both hands are used to measure the movements of their finger-joints by flex sensors, while monitoring their positions and orientations in 3-dimensional space by Polhemus sensors. The objects to be handled are observed through CrystalEYES, a pair of glasses with shutters of liquid-crystal filters. The two different images are sent separately to the right and the left eyes through the glasses, and then, a human can see 3-dimensional dynamic images.

The trainee can move pointers, small balls, in 3-dimensional space by using the dataglove system and can select and/or move objects by controlling the pointers. He can grasp or pinch the selected object by 'grasping' action through the dataglove.

Haptic feedback is important in order to manipulate objects in virtual environment. In this study, however, it is not used mainly due to hardware restriction. Sound effect is expected to compensate the lack of the haptics.

4. OBJECTS FOR EXPERIMENT OF ASSEMBLY AND DISASSEMBLY

The objects used in this study for assembly and disassembly are shown in Fig.3 which is a copy of the image on a monitor. The object to be disassembled is obtained by simplifying the check valve shown in Fig.1. The authors consider that this object meets minimum requirements for demonstrating effectiveness of the maintenance training in virtual environment.

5. GESTURE RECOGNITION ALGORITHMS

Gestures are used in two different ways: One is the case that only a pose of a hand (posture) at a moment or in a period has a meaning. Stretching a forefinger with others bent may mean to go in the direction pointed by the finger. The other is the case that a sequence of different postures has a meaning. The latter, a dynamic one, is convenient to represent various operations in the virtual environment. To grasp or release an object, to turn a handle, to turn a nut with a wrench and so on are easily expressed by dynamic gestures which are intuitively understood. In the training system in virtual environment, natural and intuitive dynamic gestures should be used so that trainees can concentrate on the training.

In Tezuka, et al. (1994) a simple posture recognition algorithm is presented. It uses a bit masking technique to remove the uncertainty of postures which differs with persons. And this idea can be used for dynamic gesture recognition.

As shown in section 2, 'select object', 'pick (grasp)', 'release', 'transfer' and 'rotate' are the important basic tasks in the system. 'Select object' is done by using two pointers which corresponds to right and left hands respectively. For example, an object is selected when a hand makes 'grasping' gesture and its pointer is near enough to the object. Judging if a pointer is near enough to an object depends on values of thresholds used in the system, which are decided by trial and error. Because the human's sense of depth is dull, the threshold value about depth should be set rather large. 'Transfer' and 'rotate' are easily realized by using Polhemus sensors. 'Grasp (pinch)' and 'release' are represented by movement of the fin-

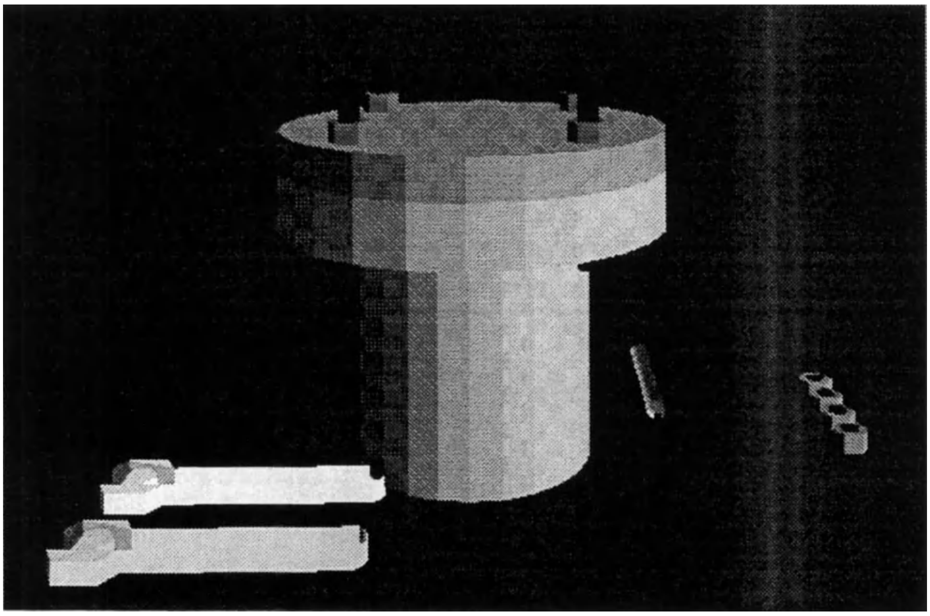


Fig. 3. Experimental objects. An object to be disassembled and a marking pen are in the center, two wrenches are in the left and four nuts are in the right.

gers. These gestures are recognized by observing the changes in the joint angles of some of fingers, and the bit masking technique is used in this recognition process.

In this system, gestures of both hands are recognized at the same time. Let's take a task, 'attaching a nut to a bolt', as an example. In this case, needed is the rule that turning a nut near the end of a bolt means to attach the nut to the bolt. The basic tasks executed on one object is not so many, usually limited to around five. Therefore, setting rules to each object or combination of objects is only a simple task.

6. TASK REPRESENTATION BY PETRI NET

In this section, the use of Petri net for representing various tasks is proposed. This representation method is explained by taking 'double-nut method' in assembly and disassembly procedure as an example.

Figure 4 shows how to pull a stud bolt out of a lid by using double-nut method. (2): A nut, nut-1, is attached to a bolt, first. (3): Then another nut, nut-2, is also attached, and the two nuts are turned by using two wrenches so that they may stick to each other. (4): Lastly, the bolt can be turned and pulled out by turning nut-1.

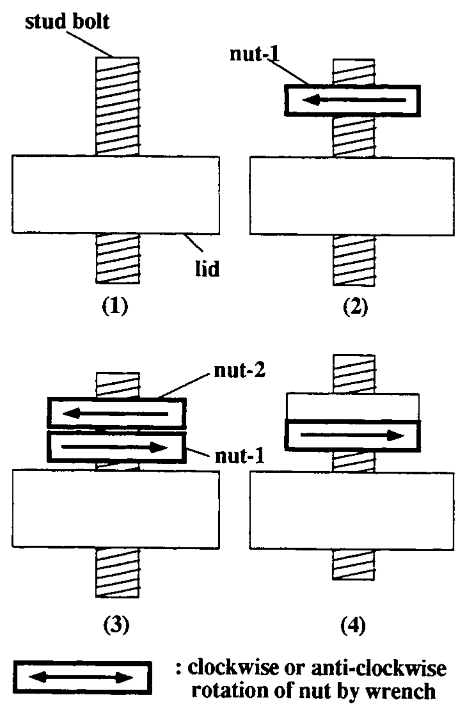


Fig. 4. Explanation of double-nut method.

This procedure can be represented by Petri net as shown in Fig.5. As the space for the paper is limited, the explanation about Petri net is omitted (Peterson, 1981). In this representation, 'place' indicated by a circle means an object (or attached objects) in some state and 'token' indicated by

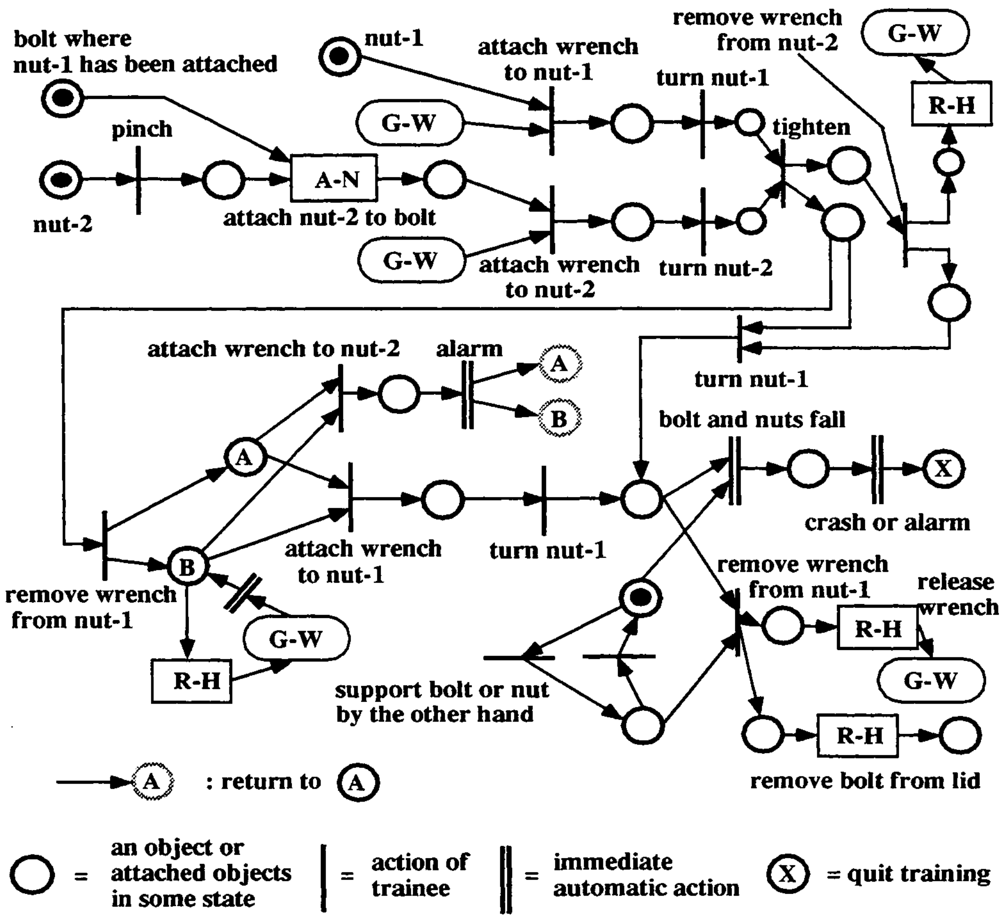


Fig. 5. Representation of double-nut method by Petri net.

a black dot in a 'place' means that the object in the state, which corresponds to the place, exists. 'Transition' indicated by one line segment means an action of a trainee, and its fire needs his decision. In this representation another type of transition is defined. 'Transition' with two line segments means an immediate action executed automatically and it fires regardless of the trainee's intention. For example, 'a wrench on a table' is represented by 'place'. 'Grasp a wrench' is represented by 'transition'. Free fall of a nut and an alarm given by a computer are represented by 'transition with two segments'. This method is convenient for examining priority of tasks or conditions for executing tasks.

In Fig.5, some tasks which appear frequently are represented by using macros for convenience. The macros are classified into place macro and transition macro (See Fig.6).

It should be noted that the trainees will not necessarily do tasks according to maintenance manuals. He may pick a nut from a table and put it on the table again, or he may hold the valve to examine its structure from various angles, whenever he wants to do so. These movements are not related to maintenance procedure, but they are not mistakes and should be treated in the system. The transition macro for pinching in Fig.6 is an example for treating this problem. This kind of macros should be prepared for all the objects which can be moved. In Fig.5 this kind of macros are omitted to simplify the net. Trainees can behave rather freely in the virtual environment owing to the existence of these macros.

The use of Petri net is beneficial for representing various tasks which have various restrictions. Another benefit of using Petri net is that the automatic assembly and disassembly can be executed easily by simulating the Petri net. Therefore, a trainee can see the maintenance procedure in the virtual space whenever he likes.

Making the Petri net for the maintenance procedure takes much time now because it is done manually. The support system for making the net should be developed, and the authors consider that it will not be so difficult a problem.

7. CONCLUDING REMARKS

In this paper, the training system developed by the authors is outlined. The gesture recognition algorithms and the representation method of various tasks by using Petri net are proposed, and

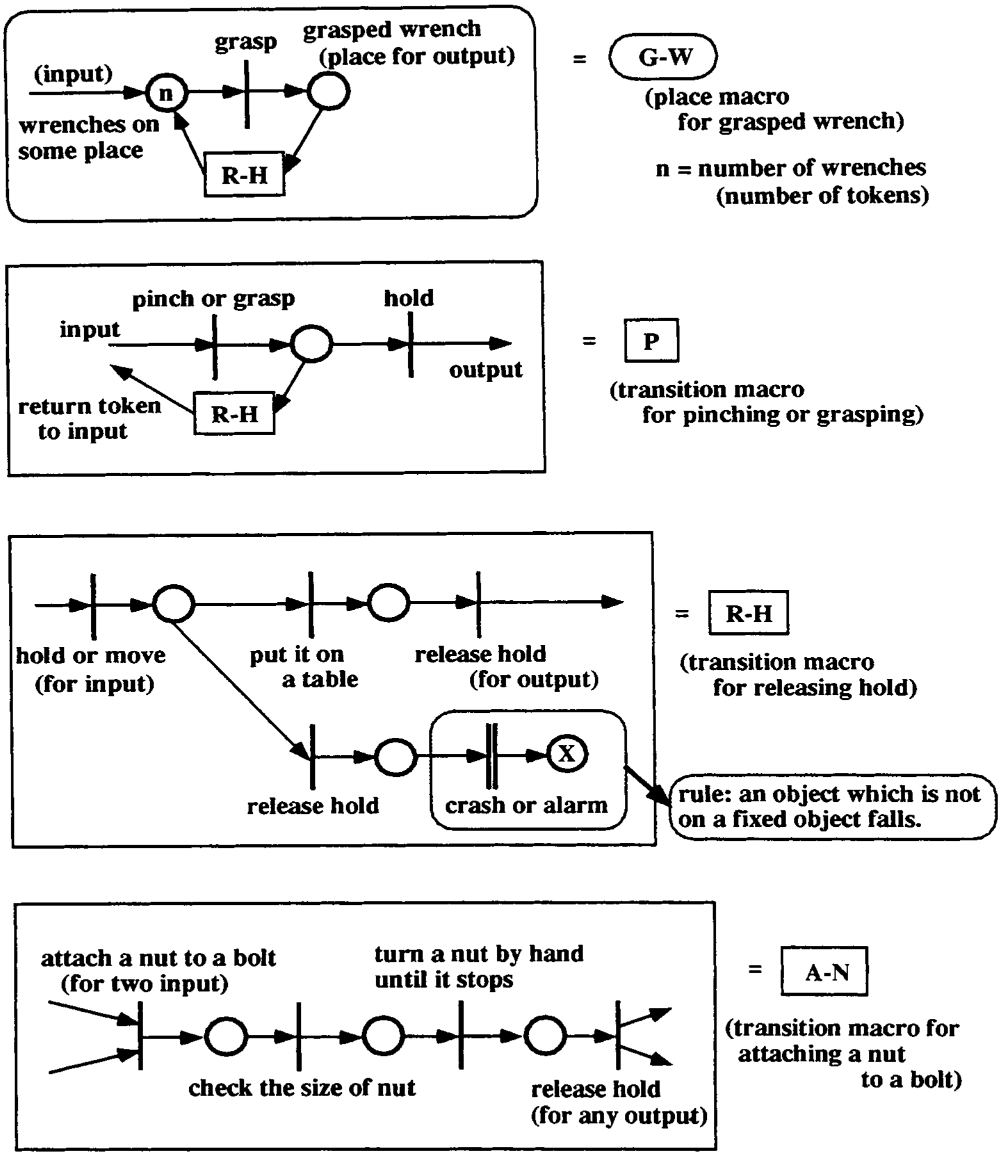


Fig. 6. Macros of tasks represented by Petri net.

their effectiveness is demonstrated through simple examples. The problem in the current system is that the way of changing viewpoint is not natural. This problem is peculiar to the system without a head-mounted display. A natural way using information obtained by the hands' movement is now under development.

Virtual reality technology is not so special as many people consider. It has a potential to prevail among various systems in industry and daily life. And its availability in future will be promising, taking the trend in the computer-technology development into account.

ACKNOWLEDGEMENT

The authors express their sincere gratitude to Mr. Ryutaro Kawano of Nuclear Power Research & Development Center, Tokyo Electric Power Company, for the research grant and his valuable suggestions.

REFERENCES

Tezuka, T., Goto, A., Kashiwa, K. and Yoshikawa, H. (1994). *A study on space interface for teleoperation system*, Proceedings of 3rd IEEE international workshop on 'Robot and Human Communication', pp.62-67.

Papper, M.J. and Gigante, M.A. (1993). *Using gestures to control a virtual arm*, Virtual Reality Systems (Earnshaw, R.A., et al. (Ed.)), pp.237-246, Academic Press Limited.

Peterson, J. L. (1981). *Petri net theory and the modelling of systems*, Prentice-Hall, Inc.

Industrial and experimental applications of transformation theory and ergodynamics

Valery F. Venda and Ilona V. Venda
Department of Mechanical and Industrial Engineering , University of Manitoba,
Winnipeg, Canada, R3T 2N2

Transformation theory and ergodynamics are used in design of RSI-free and highly productive assembly, material handling and testing workstations, analysis of control human-machine systems safety, in examining of the causes of low practical efficiency of ergonomics laboratory studies of the human-machine systems, in increasing accuracy of the performance efficiency and reaction time prediction models in comparison with traditional linear and monotonic ones.

Keywords: human factors, ergonomics, transformations, information display, cognitive science, systems safety

Introduction

One of the major goals of the current human-machine studies is increase work productivity and decrease losses caused by the repetitive strain injuries (RSI) or cumulative trauma disorders (CTD). On December 28, 1994 Investors Business Daily published a report of US Labor Department stating a 10% increase of RSI during 1994 with 302,000 workers claiming injuries versus 34,700 workers claimed RSI's in 1984. There are several detailed surveys on RSI and ergonomic studies of their prevention (Marek, Wos, Karwowski, Hamiga, 1992, Fisher, Andres, Airth, Smith, 1993). We studied experience on RSI at the assembly plants of Northern Telecom and other electronics, communication and computer companies. It was found that neck and shoulder RSI are the most widely spread traumas among the workers of the slidelines where printed circuit boards (PCB) are assembled. The RSI's are caused by bending a worker's neck along with lifting arms during a work shift at assembly slideline (Aaras, Fostervold, Thoresen, and Larsen, 1994; Schuldt, Harms-Ringdahl, Nemeth, and Arborelius, 1987).

Special attention was paid to the correlation between laboratory ergonomics experiments results and real industrial work functional structures in order to improve practical use of the human-machine systems experimental studies.

Design of new assembly, material handling and testing RSI-free workstations.

We found that all current designs of the assembly workstations were based on the assumption that it is possible to find solution optimal both, for the worker's arms and eyes. Using methodology of ergodynamics (Venda and Venda, 1991, 1995), we started to study work functional structures of the arms and eyes. At Figure 1,a acceptable angles of the assembly work surface (PCB) for the arms operations and acceptable angles of the visual object (display, screen) are shown. Position of the head (angle C) and thus the neck flexion depend on the angle of the visual object. The smaller the angle of the object, the larger flexion of the neck. At the same time position of the arms depends on the position of the PCB: the higher PCB is better observed by the eyes, but lifting the arms increases the shoulder muscle strain. The work functional structures for the arms and eyes are shown at Figure 1,b. They have a bell-shape in accordance to the First Law of Ergodynamics (Venda, 1994). In accordance to the Third Law of Ergodynamics the efficiency of interaction (joint work) of the eyes and arms, or their relative work comfort (as a probability to avoid strain injuries). As it is shown at Figure 1,b the relative work comfort (RWC) of the hands and shoulders, from one side, and of the eyes and neck, from other side, is about 0.3 of their maximal RWC values.

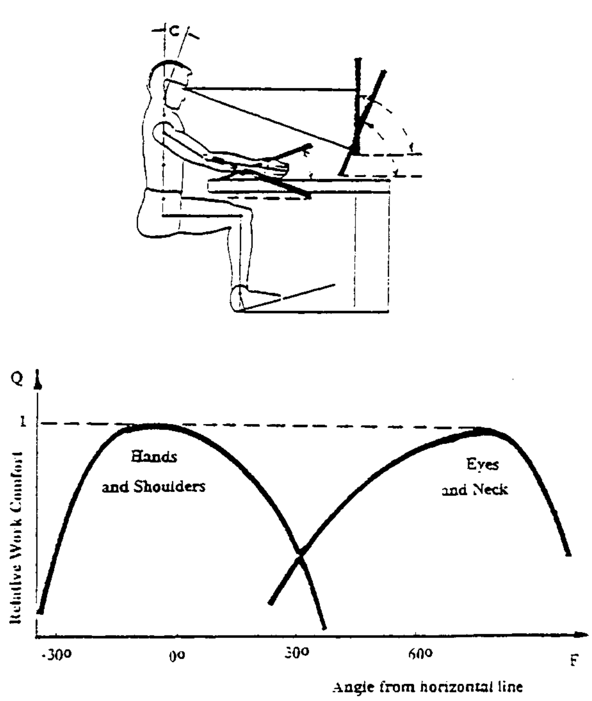


Figure 1. Work surfaces angles recommended in ergonomics literature for the eyes and arms (a), and work functional structures for the Hands-Shoulders and Eyes-Neck (b).

Therefore ergodynamics analysis showed that the traditional ergonomic approach to the design of the assembly workstations is wrong, no solutions could be found which was optimal both for the eyes-neck and for hands-shoulders at the same time. Thus we understood that the work surfaces for the arms and eyes should be separated. We designed new types of the electronics assembly workstations, the Venda-workstations (Figure 2). The workstations designs are being patented. Wide laboratory and industrial testing of the workstations was started together with Northern Telecom and Bell-Northern Research. Preliminary results are showing the Venda-workstations lower essentially risk of RSI and increase productivity and quality of some types of the assembly and material handling operations.

Our goal is to re-engineer an electronic assembly plant to improve productivity and lower the risk of RSI. In addition to the assembly and material handling workstations (the Venda-workstations) V. Venda also invented a new type of testing workstations. It is well known that testing operations take a lot of time. At some electronic telecommunication assembly plants an assembly itself takes 30% of time and testing takes almost 70% of time.

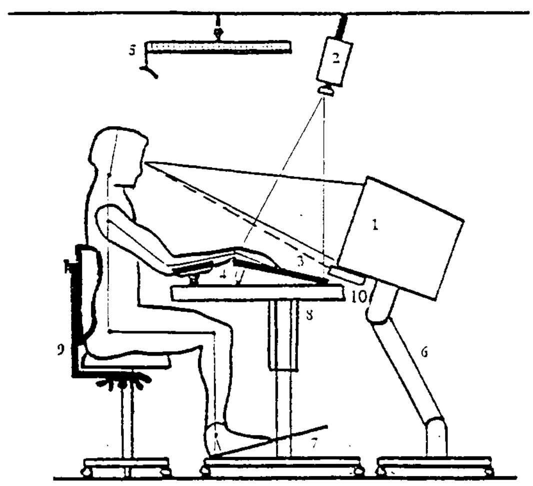


Figure 2. New type of assembly and manual material handling workstations ("Venda-assembly workstations") allows optimize positions of the Hands-Shoulders and Eyes-Neck separately. This figure illustrates also the newly invented testing workstation with hand-free interaction between tester (worker) and computer presenting testing menu, manuals, instructions. 1- screen displaying PCB and assembly (handling) operations (or computer for testing); 2 - TV camera, mirror for assembly and material handling; 3 - adjustable tilt, the workstation allows a positive, zero and negative tilt of PCB (any object handled or tested); 4 - adjustable arm rest; 5 - adjustable light; 6 - adjustable screen, TV or computer support; 7 - adjustable foot rest; 8 - adjustable desk height; 9 - adjustable chair, 10 - infrared light source and TV camera to catch tester's eye fixations for hand-free interaction with computer (only for testing workstations).

The new testing workstation, in principle, is seen at Figure 2. Infrared light source and video camera (10) catch the eye fixations of the worker at the certain parts of the test manual, menu or measurements displayed at the computer screen (1). Therefore the tester's hands are free from operations with computer and may be much more effectively used to expedite testing process. Eye Gaze System by LC Technologies was purchased to use in this new testing workstation. The click on the computer screen elements and menu positions selected may be maid by the elbow, foot, eye or head movement. Patents are pending on both the assembly and testing workstations. These are only two examples of how the ergodynamics approach helps to solve practical ergonomic problems in human-machine systems.

Ergodynamics in increasing human-machine systems safety

Considering the third law of ergodynamics (Venda, 1994), we suggested that when a normal situation in the control system is turning into an emergency one, (1) the human operator is changing a normal work functional structure into an emergency one, and (2) during the transformation period the reliability and efficiency of the human performance is lower than it was in the normal situation. To test this we ought to organize special emergency experiments at the power plant.

Using the analysis of eye movements, operators' self-reports, and detailed post-emergency interviews, we found operators used three different cognitive strategies: S_a = information perception simultaneously by small groups of elements (2-4); S_b = information perception simultaneously by middle-sized groups of elements (5-7 functionally connected elements); S_c = information perception simultaneously by big chunks (8 and more elements).

Figure 3 shows the characteristic curves Q(F) of the functional-structures (cognitive strategies) S_a, S_b and S_c on the left side and the efficiency dynamics Q(T) on the right side.

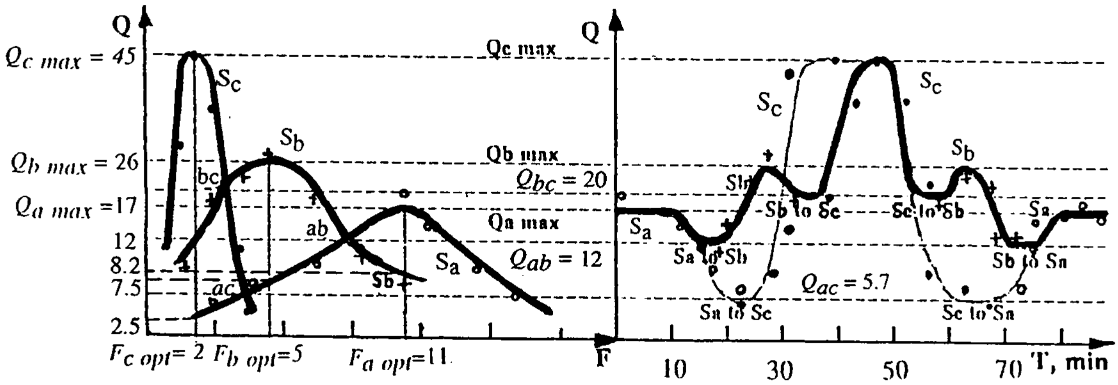


Figure 3. Work functional structures Q(F), as cognitive strategies S_a, S_b and S_c (left side) and the efficiency transformation dynamics Q(T) (right side) during emergency experiments at the power plant. Q - mean number of correct commands and control actions in ten minutes; F - mean number of eye fixations at display on every command or control action; T - time, minutes

The experiments based at organizing real emergency situations at the power plant have showed that safety of human-machine system in emergency depends on the types of the cognitive strategies used by the operator, the tempo of the transformations between the strategies and adequacy of the information display system to the actual cognitive strategies used by the operator.

Ergodynamics and practical use of human-machine laboratory studies.

The second law of ergodynamics explains why many ergonomic projects fail if they were tested at the laboratory by the subjects using one functional structure (S_{sub}) and implemented in industry where operators use another strategy (S_{op}) - see Figure 4.

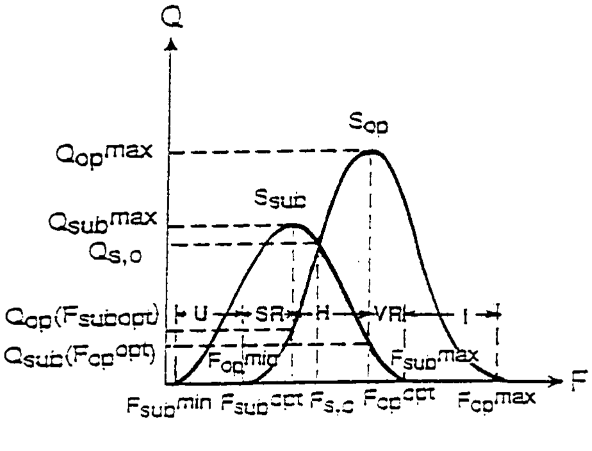


Figure 4. Work functional structures of real operators (S_{op}) and test laboratory subjects (S_{sub}). Q - criterion of efficiency (productivity); F - work factor; intervals of F: U - useless, SR - slightly relevant, H - harmful, VR - very relevant, I - inaccessible for the subjects.

There are several intervals of the values of work factor F to test the system. Testing data obtained by the subjects in the interval F_{sub}^{min} - F_{op}^{min} is useless (U), because operators do not work in this interval. Usually this is an interval of the tasks of very low complexity irrelevant for the practice but favorite at the university labs. The interval F_{op}^{min} - F_{sub}^{opt} is slightly relevant (SR), interval F_{op}^{opt} - F_{sub}^{max} is very relevant (VR) and helpful because correlation between subjects' and operators' work efficiency is positive in this interval. Thus all changes in the system design which increase subjects' efficiency in this interval will also increase operators' work efficiency. Interval between F_{sub}^{opt} - F_{op}^{opt} seems to be the most helpful because the maxima of efficiency for the laboratory test subjects and real operators are very close. But there is an unexpected paradox: in reality all data obtained in this interval F_{sub}^{opt} - F_{op}^{opt} are harmful (H). Q_{sub} and Q_{op} have a negative correlation.

Let us explain this using data of the emergency experiments at the power plant presented at Figure 3, left side. Consider a project on the ergonomic re-design of the information display for that power plant. For example the power plant managers had found the operators make only 12 correct control actions and commands per ten minutes and thus eliminate the emergencies too slowly. (Everyone knows that during emergency with quick and dangerous dynamics a slow and late decision very often is equal to the wrong decision.) The ergonomist designer-researcher gets a display that currently in industrial use and tests it at the laboratory with the subjects (university professors very often invite students to be the test subjects) using the simplest cognitive strategy S_a perceiving information elements by very small groups. Working with the display the subjects make in average nine eye fixations on one correct control action. The ergonomist tries to increase efficiency of the subjects' performance (as a number of correct control actions in ten minutes) and changes the display hardware (components set and layout) and software (particularly a tempo and volume of information presented). If new mockup of the display will force the subjects to make less eye fixations their efficiency will drop: for F=5 Q_a=7.5. Then the designer decides to change display so to increase the number of eye fixations per one action and eventually finds $F_{a\ opt}$=11. Efficiency of the subjects for their particular strategy S_a will be maximal: Q_a (F=11) = 17 = $Q_{a\ max}$.

Now consider that experienced industrial operators much more often use S_b instead of S_a. Then new display being implemented to the control room of the power plant will lead to the efficiency Q_b(F=11) = 8.2.

Therefore changing the display that **increased efficiency of the laboratory subjects** from initial Q_a(F=9)=12 to the final Q_a(F=11) = 17 and thus considered by the ergonomist as an improvement and practical achievement, at the same time actually **decreased efficiency of the real operators** from Q_b(F=9)=12 to Q_b(F=11)=8.2.

This is a concrete example how ergonomists may make harm for the industrial clients if they do not use ergodynamics and study work functional structures (cognitive strategies) used by the operators and laboratory subjects.

Therefore *the second law of ergodynamics explains why many practitioners do not trust ergonomic laboratory testing of the real systems and still prefer to relay on their experience and common sense. The paradox* is that improvement for the people with one functional structure (professional skills, individual abilities) may cause losses for the people with other functional structures. Or in other words, *the better in the laboratory the worse in reality, if the functional structures are not identical.*

Now here is an example how **ergodynamics may help to increase scientific output of experimental studies** and improve performance prediction models.

Rowe, French, Neville, and Eddy (1992) had done excellent experimental study on cognitive performance degradation during 30 hours of continuous operator performance with sleep deprivation. They collected data on the reaction time dynamics and the accuracy of operator performance. Figure 5 (left side) displays their experimental data on the performance efficiency (as an accuracy score of problem solving) dynamics presented as square shaped dots. The authors suggested approximate linear prediction model for the performance dynamics. Figure 5 (right side) shows our transformation approximation of their data. Figure 6 (left side) displays their experimental data on reaction time dynamics as square shaped dots and their theoretical linear prediction as a solid about monotonic line. Figure 6 (right side) shows our transformation approximation of their data on reaction time dynamics.

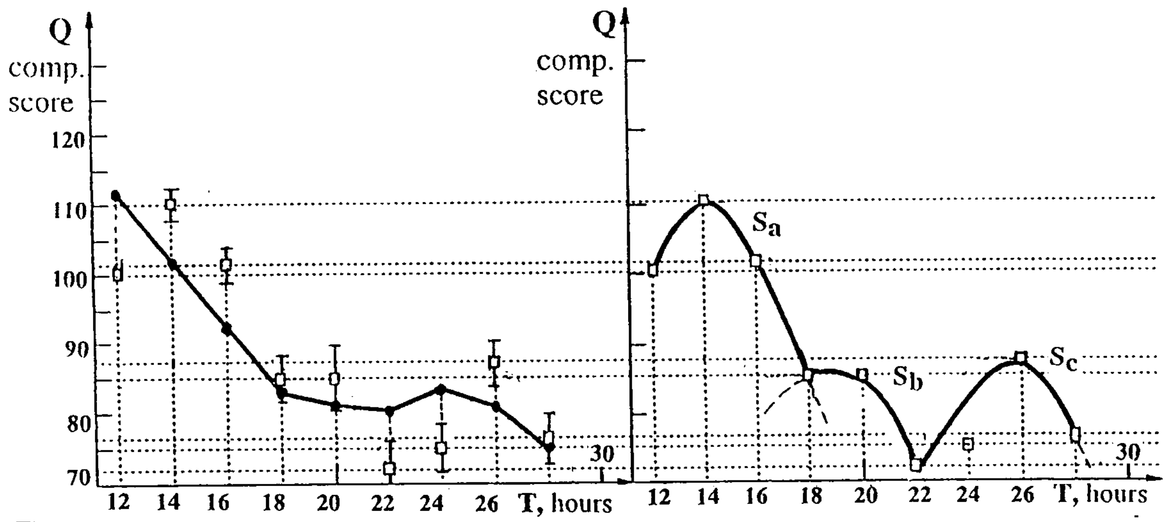


Figure 5. Experimental data by Rowe et al. (1992) on the performance efficiency dynamics (dots) and the authors' predicting model (solid line) on the left side; and my approximation of their data and performance predicting model based on the transformation theory and ergodynamics on the right side. Q - performance efficiency as an accuracy score, T - time of the experiment duration, hours.

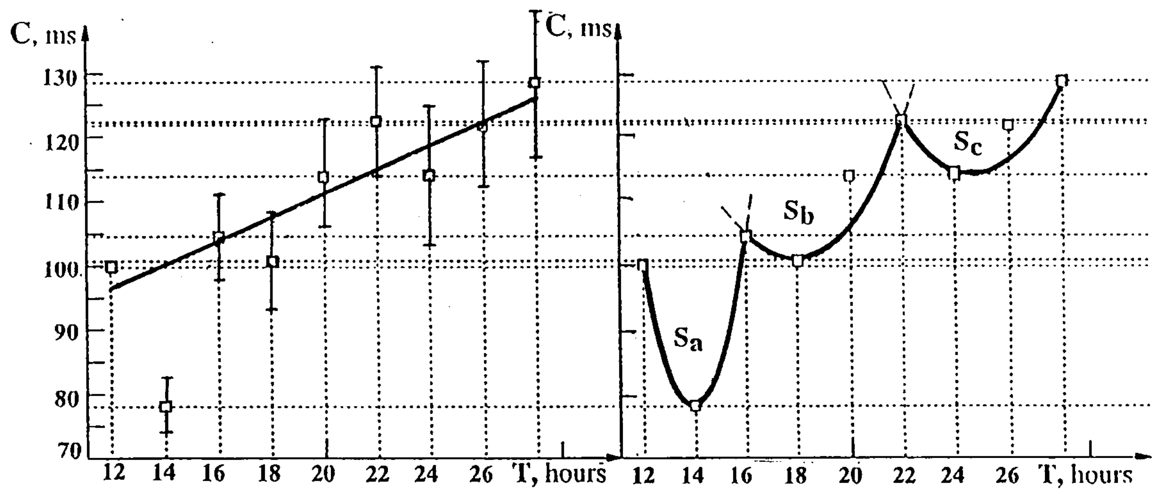


Figure 6. Experimental data by Rowe et al. (1992) on the reaction time dynamics (dots) and the authors' predicting model (solid line) on the left side; and our approximation of their data and performance predicting model based on the transformation theory and ergodynamics on the right side. C - reaction time, ms; T - time of the experiment duration, hours.

Our analysis of the data by Rowe et al. (1992) on the efficiency dynamics (Figure 5) and the reaction time dynamics (Figure 6) suggested there were three different functional structures of human performance used and transformed during 30 hour experiments. On the right sides of Figure 5 and 6, we show hypothetical characteristic curves of the work functional structures S_a, S_b and S_c. The mean value of deviations between experimental data by Rowe et al. (1992) and their prediction model for efficiency dynamics is $M_{Q\ Rowe}=5.77$; for the reaction time it is M_{CRowe}=6.33. Our prediction model, based on transformation theory, gives the following respective numbers: M_{QVenda}=1.15; M_{CVenda}=1.55. Therefore the accuracy of our transformation models for this performance dynamics prediction is higher than those traditional linear and monotonic models used by Rowe et al. (1992) by $5.77/1.15 \approx 5$ times for the efficiency prediction and $6.33/1.55 \approx 4$ times for the reaction time prediction.

How did we prove that the three work functional-structures were successively used at the experiments by Rowe et al. (1992)? The work functional structure is an integral unit with synchronic changes of complexity and efficiency (Venda and Venda, 1995). The experiments have shown that the complexity criterion (average normalized reaction time score, C) and efficiency criterion (average normalized performance accuracy score, Q) had the extreme values (minimums of complexity criterion, C, and maximums of efficiency criterion, Q) exactly at the same time periods: the 14th hour, 18th hour, and 25th hour.

Conclusion

Transformation theory and ergodynamics are very helpful in solution of practical problems on design and improvement of human-machine systems, RSI free and highly productive workstation for assembly, testing, inspection, material handling operations, in increase of industrial applications of the ergonomics laboratory studies, in achieving higher safety of the control human-machine systems.

Acknowledgments

This studies are being supported by Northern Telecom Canada, Bell-Northern Research and Natural Science and Engineering Research Council of Canada.

References

Aaras, A., Fostervold, K. I., Thoresen, M., and Larsen S. (1994) Postural load at VDU work: A comparison between different workplace designs. Personal communication.

Fisher, D.L., Andres, R. O., Airth, D., and Smith, S. S. (1993) Repetitive Motion Disorders: The Design of Optimal Rate-Rest Profiles, Human Factors, 35(2), 283-304.

Marek, T., Wos, H., Karwowski, W., and Hamiga, K. (1992) Muscular Loading and Subjective Ratings of Muscular Tension by Novices when Typing with Standard and Split-Design Keyboards, International Journal of Human-Computer Interaction, v. 4, #4,387-394.

Lomov, B. F., Venda, V. F. (1977). Human factors; problems of adapting systems for the interaction to the individual: The theory of hybrid intelligence. Proceedings of the Human Factors Society 21st Annual Meeting, San Francisco, 1-9.

Rowe, A. L., French, J., Neville, K.J., and D. R. Eddy (1992) The prediction of cognitive performance degradations during sustained operations, In: Proceedings of the Human Factors Society 36th Annual Meeting, Atlanta, GA: HFS, 111-115.

Schuldt, K., Ekholm, J., Harms-Ringdahl, K., Nemeth, G., and Arborelius U. P. (1986) Effects of changes in sitting work posture on static neck and shoulder muscle activity. Ergonomics, Vol. 29, No 12, 1525-1537.

Venda, V. F., and Venda, Yuri V. (1991) Transformation dynamics in complex systems, Journal of Washington Academy of Science, #4, December.

Venda, V. F. (1994) Ergonomics: theory and tips. Manifesto of ergodynamics - Keynote Address for the World Ergonomics Congress IEA'94, Toronto, August, 1994. Proceedings of the 12th Triennial Congress of International Ergonomics Association IEA'94, HFAC, Toronto, 34-36.

Venda, Valery F. and Venda, Yuri V. (1995) Dynamics in Ergonomics , Psychology, and Decisions, (Ablex, Norwood, in press).

Transformation dynamics in human-machine systems

Valery F. Venda
Department of Mechanical and Industrial Engineering , University of Manitoba,
Winnipeg, Canada, R3T 2N2

Abstract
Transformation dynamics theory and its applications to human-machine systems and ergodynamics, are based on three laws: 1) mutual adaptation, 2) plurality of work functional structures, and 3) transformations. Experiments to discover, prove and illustrate the laws were conducted. Ergodynamics aids in studying the transformations that occur in human operator work functional structures, skills, and cognitive strategies as transitions in the structures of processes of human-machine-environment mutual adaptation. Using principles of ergodynamics in human operators training increases their ability to transform quickly a control strategy adequate in normal situation into another one adequate in emergencies, therefore improving reliability and safety of human-machine systems.

Keywords: human factors, ergonomics, human-machine interface, cognitive system, transformations.

Introduction

Since the Ebbinghaus' (1885) studies on learning, the monotonic learning theory was the only one used to describe the dynamics of human performance efficiency versus time of training. The monotonic curve has thus gained ground in applied psychology and ergonomics. The main reasons that led to the "monotonic monopoly" in experimental and mathematical psychology were: (1) massive data on experiments with humans and animals performing simple tasks using a single strategy and (2) the theory allowed for simple and convenient mathematical exponential models. Ergonomics created one more reason for the expansion of the application of Fitts' law beyond its inner margins (Fitts, 1964, Sheridan, 1992). Fitts (1964) limited this law only to the tasks of reaching for controls and tools by hand, given a certain initial position (Konz, 1990). See Sheridan (1992) for a greater explanation of these limitations.

In *Introduction to Ergonomics,* Zarakovsky et al. (1977) (based on an interpretation of Fitts' law), assumed that *monotonic dependence of operator work efficiency (Q) on information volume perceived (I) automatically leads to a monotonic learning process: dynamics of Q in time (T) (see Figure 1).* Their proposed curve tends towards one asymptotic "maximal" level of efficiency, speed, skill, etc..

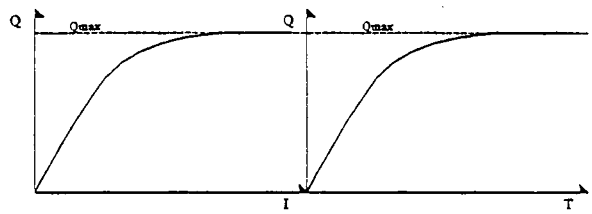


Figure 1. Assumption on monotonic dependence of operator work efficiency (Q) on information volume perceived (I)-- left side-- led to the conclusion on monotonic learning process: dynamics of Q in time (T) (Zarakovski et al., 1977)

Transformation learning theory was first published in 1980 (Venda, 1980). In 1982, a seminar on the theory was presented at MIT and Harvard University where previous, monotonic learning theory was studied intensively (Estes, 1950, Luce, 1959). There was a very interesting and very intensive discussion at Harvard following which W. Estes and D. Luce helped to edit and publish a paper on the new theory (Venda, 1986). At MIT T. B. Sheridan advised me to organize different experiments obviously showing the processes of transformations and find fundamental axioms of transformation dynamics in human-machine systems (Yufik, Sheridan, and Venda, 1992, Yufik, Sheridan, and Venda, 1993).

Simple experiments on influence of the desk height on typing productivity were organized (Venda, 1994). In parallel, long term and short term transformations in work functional (e.g. cognitive) structures were studied when human operators at a plant changed their performance from normal to emergency situations. It was found that human performance efficiency declines inevitably during the transformations of the work functional structure. Transformations of the cognitive structures were found to be a contributing factor among causes of many technological, aviation and various other accidents, including Chernobyl, Three Mile Island, and Bhopal (Venda, 1990).

Basic ergodynamics definitions.

1. *Work function* is a human goal in dynamic human-machine interaction, like creating software, producing cars, making control decisions at nuclear power plant, etc.. 2. *Work efficiency*, Q, is any quantitative measure (criterion) of a positive outcome of the dynamic human-machine interaction. It can be a number of the quality products produced, a number of correct operation or decisions made, probability of safe and no-injury work during a certain time period, etc. Q may have an actual, real value (Q_{act}); Q may have a maximal, theoretical value (Q_{max}). Q may be measured in absolute units (Q_{abs}) or have a relative value ($Q_{rel}=Q_{act}/Q_{max}$) 3. *Work complexity*, C is an opposite criterion to Q. To accomplish the work function (i.e. to reach a maximum efficiency, Q), individuals must overcome various obstacles, challenges, and difficulties. Such obstacles lower the real work efficiency value in comparison with the theoretically maximal one. We name this loss of work efficiency, resulting from these obstacles (subjective, objective, human-machine non-optimal interaction, etc.), *work complexity (C)*. For example, let us consider that the work efficiency (Q) of some facility (human-machine system) is measured as the number of products of high quality. If we assume that the total number of products produced, if they were all of a high quality, is the highest desirable efficiency level (Q_{max}), then the complexity is expressed as $C_{act} = Q_{max} - Q_{act}$. Thus we have, in absolute values, $Q_{max} = Q + C$. Alternatively, we may express the same in relative values, if dividing all these numbers results in Q_{max}. We now have $Q_{max}/Q_{max}= Q/Q_{max} + C/Q_{max}$; $Q_{rel} + C_{rel} = 1$ or simply $Q + C = 1$. When I explained this at the seminar given at Riso Lab in Copenhagen, J. Rasmussen questioned if I am adding oranges to elephants. No, I suggest that *both a positive work criterion, which is efficiency (Q), and a negative work criterion, complexity (C) are expressed in the same units of measurement* (e.g. like the number of the quality products to be sold (Q_i) and the rest of the products to be repaired or disposed of as a trash (C_i) or a probability of a correct decision (Q_j) and probability of a wrong decision (C_j)). Opposite each to other character of Q and C could be expressed also as $Q/C = 1$. For example, if C is time spent to produce one item (or to process one signal, solve one task) then 1/C will show a productivity in a time unit that is important criterion of efficiency, Q. Q and C could be defined and measured only if one knows exactly the people's work goals, functions, conditions. 4. *Work factor* F is a parameter of people-objects-environment interaction and mutual adaptation influencing work efficiency. The examples of F are the computer desk height, computer screen size, weight of hammer, resistance of the car break pedal, etc. 5. We define a function $Q_i(F)$ as *a work functional structure:* $S_i = Q_i(F)$.

6. Applied goal of the ergodynamics is optimizing Q and C. Usually, in industrial practice that means increasing Q and decreasing C (Venda, 1994), but in education, science the goals may be different (Savelyev and Venda, 1989). The equation $Q + C = Q_{max}$ leads us to three major strategies to increase the work efficiency (Q): 1) *extensive strategy*: increasing Q_{max}, and thus increasing Q, when Q/C = Cons.; this strategy requires extra resources and expenditures, including human energy and time; extensive strategy actually changes the content of the system, and thus the system itself; a company should hire more people, buy more equipment and efficiency will grow, but complexity or losses will also grow; this is a strategy of additional investments and losses; 2) *intensive strategy*: making an accent on the efficient, high quality work, expanding its share in the work shift, decreasing time delays and unnecessary breaks; in this case, people should work harder and management should re-train workers to improve their attention and work skills; this is the strategy of harder work; 3) *strategy of mutual adaptation and transformation*: in this case, one makes an accent on decreasing complexity, eliminating the obstacles or causes of errors; this strategy is based on a better mutual human-machine-environment adaptation; this is the strategy of smarter work and design.

7. Optimizing work efficiency is always a process of human-machine-environment mutual adaptation: people improve their skills and knowledge, change the objects and environment to harmonize whole interaction system. My 30 years of experience in

studying dynamic human-machine-environment interaction led me to the following conclusion: in order *to reach maximal work efficiency one should organize mutual adaptation between human work functional structure* (by professional selection, training, motivation) *and work environment* (by ergonomic design of machine, interior, software) (Venda, 1969, 1975, Venda and Venda, 1995).

The First Law of Ergodynamics.

We conducted the following experiment. Our subjects were typing on the notebook computer in sitting position at different desk heights. Typing productivity vs. desk height obtained at the experiments is shown at Figure 2. It has a bell-shaped function. There are many literature data confirming this regularity.
The First Law of Ergodynamics (The Law of Mutual Adaptation) reads as follows: "Work efficiency is a bell-shaped function of the factor of mutual adaptation between human work functional structure and work environment."

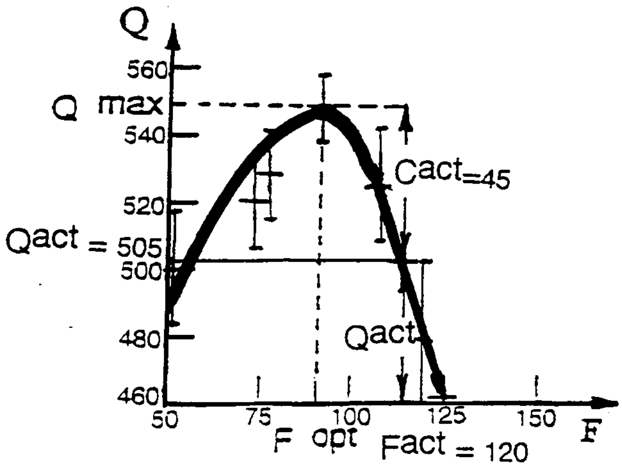


Figure 2. Efficiency of typing, Q, characters/min., versus desk height, F, cm.

Figure 2 presents a simple visualization of the First Law of Ergodynamics. There are two ordinates: efficiency, Q, and the factor of human-environment (machine, other person) mutual adaptation, F.
To use the First Law of Dynamic Interaction, or Ergodynamics in practice of Cognitive Technology one should (1) find a main factor of human-environment mutual adaptation; (2) find the direction to change the factor to increase efficiency; (3,a) if the current factor and efficiency values are at the left side of the bell-shaped curve of the work functional structure, the factor value should be increased to increase the efficiency; (3,b) if they are on the right side of the curve then the factor value should be decreased to increase the work efficiency. The First Law explains the big jump over F_{opt} brings the ergonomist to another branch of the curve at Figure 1 with an opposite Q(F) correlation. If both changes of the factor to the left and to the right led to higher efficiency criterion values then the ergonomists should (4) study two different possible work structures. Thus we found a plurality of the functional structures (Venda, 1980).

The Second Law of Ergodynamics.

The experiment on typing at the notebook computer was extended so the subjects were typing in sitting (S_1) and standing (S_2) positions with different desk heights. Two different functional structures for every subject were found (Figure 3). Earlier we studied the processes of reading, perception of control board information and found the families of bell-shaped functional structures (Venda, 1986, Venda, Stishkovskaya, et al., 1994).

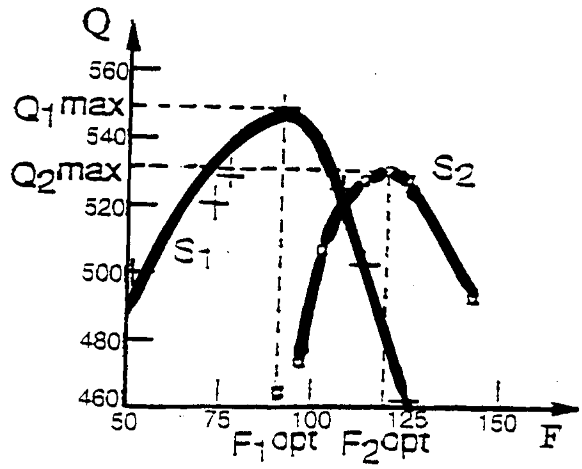


Figure 3. Functional structures of typing in sitting (S_1) and standing (S_2) positions. Q, productivity of typing, characters/3 min, F, a desk height, cm.

These studies allowed us to word the Second Law of Ergodynamics (*The Law of Work Structures Plurality*): "Every work task can be done with different work structures".
A visual image of the Second Law is shown at Figure 3 where S_1 and S_2 are two different work functional structures with optimal F values: F_{1opt} and F_{2opt}, and maximal levels of efficiency: Q_{1max} and Q_{2max}. The Second Law explains plurality of human reactions on the same signal (environment factor F). There is a crucial practical challenge: transfer of laboratory data to ergonomic design. Let us assume we compared work

efficiency Q of the laboratory subjects and real operators using the task with factor $F=F_{1,2}$. The efficiencies were equal $Q=Q_{1,2}$. But if the subjects use S_1 and real operators use S_2 then optimizing F for the subjects will catastrophically decrease efficiency of the real operators: $Q_2(F_{1opt}) \rightarrow 0$.

The Third Law of Ergodynamics.

World science (engineering, ergonomics, economics and psychology included) knows only the following dynamics of development, progress, learning: 1) step function (efficiency increases overnight); 2) linear increasing of efficiency; 3) monotonic exponential increasing of efficiency (Ebbinghaus, 1885), some times with an intermediate plateau (Brian and Harter, 1896). Most authors insist on monotonic dynamics (Chapanis, 1959).

The typing experiment was modified so the desk height was increased by intervals of 5 cm from very low (50 cm) to very high (140 cm). The subjects were typing at first in sitting position then changed position to the standing one. Overall, integral typing productivity, Q_{int} at all heights was measured. The task was to find a height to change sitting position to standing one, thus resulting in the Q_{int} being maximal. It was found that changing (transformation) of one position to another should be done at the height corresponding to the intersect point for the work functional structures in sitting and standing positions (Figure 4).

The Third Law (*The Law of Transformations*) was worded as follows:

"Transformations between different structures of the system and interaction between different systems' structures are maximally effective if they go through a state common and equal for the structures."

The Third Law explains a wave-like dynamics of the transformations of the work functional structures we discovered when conducted the following training experiment. Our subjects (twelve students of engineering faculty) were asked to accomplish a compensatory tracking of dynamic signals presented simultaneously on several (from one to six) measurement instruments with the subject controlling an equal number of switches (from one to six). Different cognitive and sensory-motor strategies and transformations between them were found (Venda, 1986). Most learning curves were non-monotonic in nature (Figure 5).

When we studied the process of signals tracking and different numbers of the signals were presented (Figure 5), traditional statistical averaging was avoided because it leads to the monotonic learning curve masking real processes of transformations (Figure 6).

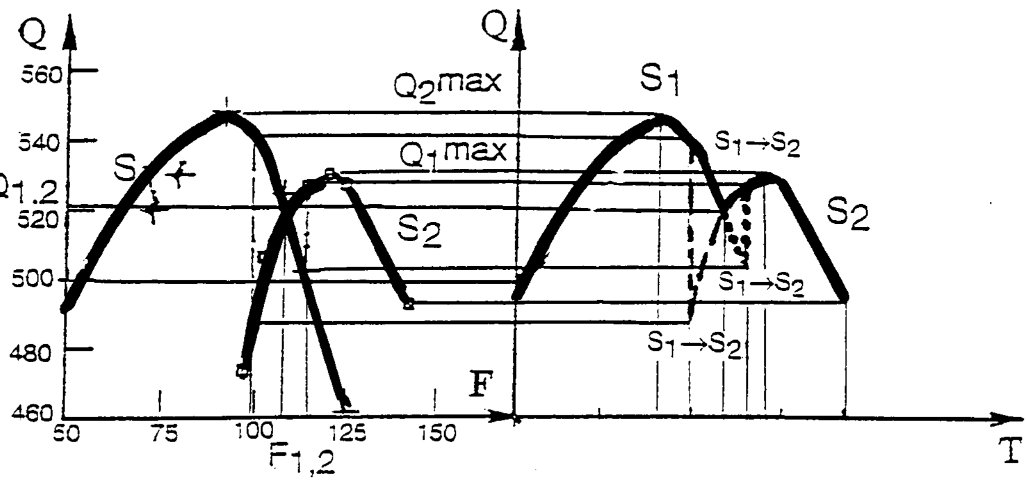


Figure 4. Comparison of the integral productivity values, Q, when transformation from sitting position to the standing one is done at different desk heights, F. If F changes gradually from F_{min} = 50 cm to F_{max} = 140 cm, transformations betweetwo functional structures S_1 and S_2,($S_1 \rightarrow S_2$), which go through a common and equal state with the coordinates $Q_{1,2}$,= 522 char./3 min. $F_{1,2}$.=107 cm at the trial #12, lead to the bigger integral productivity (as a sum of Q_i at all heights, F) than transformations at any other state, e.g. at F=100 (trial #10) and F=120 cm (trial #13) as shown with broken lines. T = time (number of trials, consecutively from F = 50 cm to F = 140 cm with difference between next to each other trials ΔF = 5 cm).

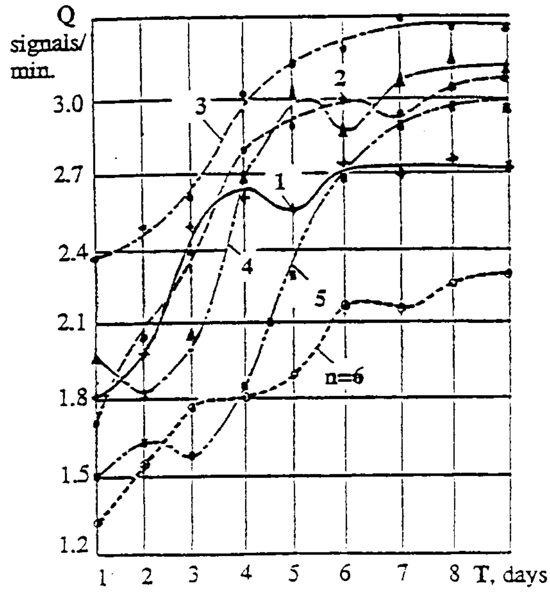


Figure 5. Dynamics of efficiency of the compensatory tracking Q (signals/minute) as the function of number of measurement instruments presented (n=1-6) and time, T (number of training sessions) - a). Statistical averaging all data shown at a) leads to the monotonic curve Q(T) - b).

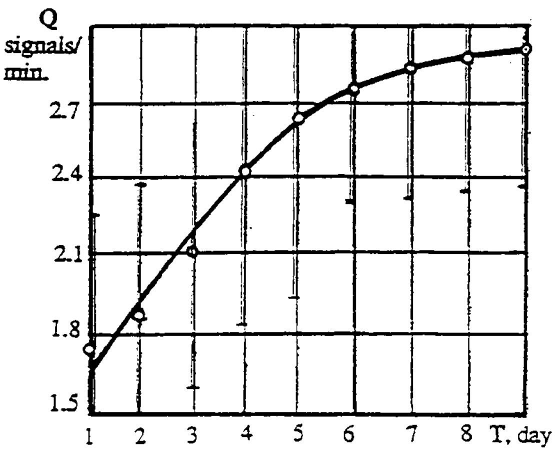


Figure 6. Averaging experimental data can mask real transformation processes in training dynamics.

A dependence of signal tracking efficiency on different cognitive strategies was then analyzed: perception of the signals by one, two and three as an information chunk. An example of this analysis, for the situation when six signals were displayed simultaneously, is shown at Figure 7. Using eye movement registration, along with an analysis of the numbers of operations in a minute, and signals tracked in a minute, we found that subjects in the beginning of training perceived the signals separately, one by one (strategy S_1), then perceived the signals by the groups of three, as the coordinates of a 3-dimensional space (S_2). If they tried to start from S_2, while their skills did not allow them to make more than 20 control operations in a minute, their efficiency (Q) was lower than using S_1. If some participants simply tried to increase the tempo of the operations, they made more operations than optimum for S_1: $F>F_{1opt}$ = 16 operations/minute, and they did not transform S_1 into S_2, thus decreasing their efficiency (in comparison with Q_{1max} = 17 signals/min).

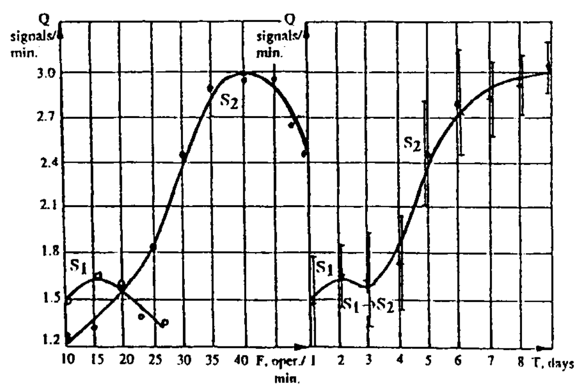


Figure 7. Analysis of different cognitive strategies as statics (on the left side) and their transformations as training dynamics (on the right side), when six signals were displayed simultaneously (n=6). Q - work efficiency as a number of the signals tracked in a minute, F - number of operations to track one signal, T - training time (numbers of training sessions), S_1 and S_2 - work functional structures (cognitive strategies): S_1 - perception and tracking the signal one by one, S_2 - perception and tracking by groups of three signals (3-D spaces).

Conclusion

Ergodynamics helps predict dynamics of transformations in human-machine systems, and helps plan the transformations to minimize losses.

Work structures may belong to the same individual and may be used sequentially, transforming one to another. This human ability gives human-machine system a very important feature of *successive flexibility*. In addition to the traditional training of the operators separately in solving normal tasks and emergency tasks, ergodynamics insists at the necessity to train human operators in quick and reliable transformations normal work structures into emergency ones and vice versa. This will significantly increase human-machine systems safety.

Work structures may belong to different individuals (or human and machine) interacting in the work processes. This system's ability gives the system another very important feature, a *simultaneous plurality*. Different functional structures of decision making may be used by different specialists of the group for the solving the most urgent and difficult tasks. This group may be organized on the principles of the hybrid intelligence systems suggested long ago (Lomov and Venda, 1977), but for some reasons are being implemented to the research and practice very slowly. Hopefully ergodynamics will help to study sophisticated problems of the hybrid intelligence human-machine systems. Detailed introduction to the theory and practical applications of ergodynamics and hybrid intelligence are presented in (Venda and Venda, in press, planned for 1995).

Acknowledgments

This studies are being supported by Northern Telecom Canada, Bell-Northern Research and Natural Science and Engineering Research Council of Canada.

My special personal thanks go to Dr. T. B. Sheridan for his generous help, discussions and suggestions during almost twenty years of our collaboration, friendship and meetings in Cambridge, MA, Moscow and many other places.

References

Atkinson, R.C., and Croters, E.J. (1964) A comparison of paired associate learning models having different acquisition and retention axioms. Journal of Mathematical Psychology, 1, 285-315.

Bryan, W.L., and Harter, N. (1899) Studies on the telegraphic language: The acquisition of a hierarchy of habits. Psychological Review, 57, 94-107.

Bush, R.R., and Mosteller, F. (1955) Stochastic models for learning. New York: Wiley.

Chapanis, A. (1959) Research techniques in human engineering. Baltimore: Johns Hopkins University Press.

Ebbinghaus, H. (1885) Uber das Gedachtnis. Berlin: Leip-Dunker.

Estes, W. K. (1950) Toward a statistical theory of learning. *Psychological Review*, 57, 97-107.

Fitts, P.M. (1964). Perceptual-motor skill learning. In:

Konz, S. (1990). Work design: industrial ergonomics, Worthington: Publishing Horizons.

Lomov, B. F., Venda, V. F. (1977). Human factors; problems of adapting systems for the interaction to the individual: The theory of hybrid intelligence. Proceedings of the Human Factors Society 21st Annual Meeting, San Francisco, 1-9.

Luce, R.D. (1959) Individual choice behavior: A theoretical analysis. New York: Wiley.

Savelyev, A. Y. and Venda, V. F. (1989) Higher Education and Computerization, Moscow: Progress.

Sheridan, T. B. (1992). Telerobotics, automation and human supervisory control. Cambridge, Mass.: MIT Press.

Taylor, F. W. (1971). The Principles of Scientific Management, N. Y.: Harper and Row.

Thorndike, E.L. (1898) Animal intelligence: An experimental study of the associative processes in animals. Psychological Monogr., #2(8).

Venda, V., (1986). On transformation learning theory, Behavioral Science, v-31, No. 1, pp. 1-11.

Venda, V.F. (1992) Transformation Dynamics in Humans, Industry, Society (Russian, American, and Canadian Experience). Keynote Address for 36th Annual Meeting of HFES, Atlanta, October 1992.

Venda, V. F., and Venda, Yuri V. (1991) Transformation dynamics in complex systems, Journal of Washington Academy of Science, #4, December.

Venda, V. F. and Venda, Y. V. (in press) Dynamics in Ergonomics, Psychology, and Decisions, Norwood: Ablex.

Woodworth, R. (1938) Experimental Psychology, New York: Holt.

Yanch, E. (1974) Forecasting of scientific and technological progress. Moscow: Progress.

Yufik, Y. M., Sheridan, T. B. and Venda, V. F. (1992) Knowledge Measurement and Cybernetics of Mutual Adaptation in Human-Machine Systems, In: Cybernetics and Applied Systems, Marcell Dekker, N. Y. pp. 187-237.

Yufik, Y.M., Sheridan, T.B., and Venda, V.F. (1993) Quantitative evaluation and performance prediction, In: G.Salvendi and M.J. Smith (Eds) Human - Computer Interaction, 19A, Amsterdam: Elsevier, 642-647.

Zarakovski, G. M., Korolyov, B.M., Medvedev, V.I., and Shlaen, P.Y. (1977) Introduction to ergonomics. Moscow: Sovetskoe Radio.

A NEW MACHINE LEARNING METHOD INSPIRED BY HUMAN LEARNING

B. YANG and H. ASADA

Massachusetts Institute of Technology, Center for Information-Driven Mechanical Systems, Department of Mechanical Engineering, Cambridge, Massachusetts, USA

Abstract. When people learn new tasks, they are slow and meticulous in the beginning but speed up the operations as they gain experience and skills. In this paper, we will explore a new approach to learning, inspired by this human learning behavior. The new method termed "Progressive Learning" uses scheduled excitation inputs that allow the system to learn quasi-static, slow modes in the beginning, followed by the learning of faster modes. This new method is presented in the context of high speed robotic assembly, where an impedance control law is learned with this excitation scheduling method. Extensive simulation results are provided to demonstrate the effectiveness of this method. A detailed analysis of the mechanism of progressive learning is also provided and verified through simulation.

Key Words. Human Learning, Machine Learning, Adaptive Control, Impedance Control, Persistent Excitation

1. INTRODUCTION

Human learning is by far more effective than machine learning in many ways. Humans use a curriculum, an organized set of materials to be learned by learners in a particular sequence. They, for example, attempt simple tasks at the beginning in order to become familiar with the tasks and then, based on the knowledge and experience gained, execute the full-scale tasks. By doing so, they can avoid costly failures and damage, or confusion in which the learner could gain neither useful information nor valuable experience. Such confusing results and failures would have no pedagogical value for learning the task. Depending on the competence and amount of knowledge the learners possess at the beginning, judicious choices are required to determine the level of task complexity appropriate for initial attempts.

Inspired by this human learning behavior, we will explore a new approach to machine learning. The level of task complexity will be increased gradually as the learner gains knowledge and improves task performance. To this end, we need to organize the learning process by evaluating the learner and determining the level of task complexity appropriate for the learner, so that the learner can attain useful knowledge with a minimum of confusion and failure. Figure 1 shows a conceptual diagram of this learning system consisting of a learning organizer, called a "pedagogue", and a learner. The learner, who has its own learning mechanism, is provided by the pedagogue with a series of programmed tasks, called a "curriculum". Designing this learning system is thus two-fold: the design of the learning rule for the learner and the design of tasks, i.e. the curriculum.

This task design problem is relatively new to the learning/adaptive control community, but relevant issues have been addressed in a different context. In adaptive control, for example, the design of reference signals has been a central issue for the convergence of system parameters. In particular, the concept of persistent excitation (PE) has played an important role in the design of reference inputs. Namely, the PE provides the necessary conditions for the parameters of a plant or a controller to converge to their desired values [8]. There are a number of papers addressing PE conditions [2] [3] [5] [7]. While these PE conditions provide fundamental conditions for stability and convergence, we need an effective "strategy" for accomplishing the conditions. To this end, we need to take a further step from the analysis of PE conditions on reference inputs to the total design and synthesis of "tasks" to be performed by the learner. Namely, a strategy must be set up that allows the system to be excited effectively in order to expedite the learning and adaptation process, while task performance requirements are being met during the process.

The objective of this paper is to explore a novel learning method in which the level of task complexity advances progressively in accordance with

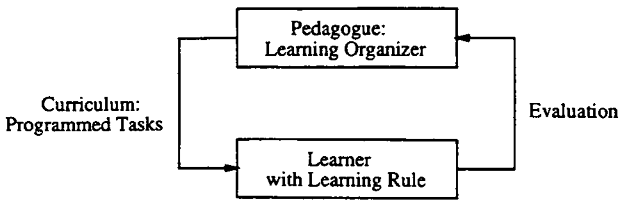


Fig. 1. Conceptual diagram of progressive learning system

the learner's competence and level of accomplishment. We design a series of tasks with different complexity levels appropriate for the learner. The idea is to integrate the task assignment scheduling with the design of the total control algorithm. It is expected that by integrating these we will be able to avoid instability and divergence of learning, expedite the learning process, and maintain a desired task performance level having a minimum chance of failure and damage to the system.

2. LEARNING APPROACH INSPIRED BY HUMAN

To apply on-line learning methods to practical processes, stability must be guaranteed. Learning algorithms that cannot be guaranteed to converge within a certain time limit are not acceptable or feasible for practical use. Moreover, in most applications, a certain minimum level of task performance must be accomplished at all times, even at an early stage of the learning process. Once the system is engaged in an actual task, it must not fail in performing the task, nor yield poor outputs. Learning algorithms, although guaranteed to converge in theory, may not be applicable to real-time, on-line learning if task performance during the learning process is not satisfactory. In particular, the early stage of learning when the system does not have enough data or exact knowledge about the task is difficult to deal with.

Humans perform unfamiliar tasks slowly and meticulously when their knowledge is limited and stringent task specifications must be met. In the early stages of learning, we first lower the task complexity in order to avoid failure and unsatisfactory performance, and then increase the complexity level in accordance with the progress in learning and the improvement in task performance. By reducing speed, for example, the task is made tractable and executable despite limited knowledge and skills. As humans gain experience and become familiar with the new task, they increase the task execution speed. The task complexity can also be lowered by relaxing some task conditions and limiting the scope of the task. Learners can deal with tougher conditions and a broader range of situations after they have learned the simplified tasks.

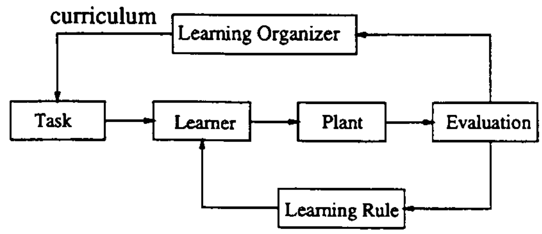


Fig. 2. Progressive learning system

If this strategy is successfully applied to machine learning, we will be able to overcome the difficulties of traditional learning methods as described above. In order to learn a task while executing it by satisfying minimum task requirements, learning and task execution must be coordinated. In the traditional framework of learning control, the major research interest is focused on the development of learning rules and their convergence conditions. In the proposed approach, we do not address a learning rule alone, but we integrate it with the design of a series of scheduled tasks in which the task complexity level is varied depending on the learning progress. This new learning method, which we refer to as "Progressive Learning", is defined as follows:

> Progressive Learning is a learning method in which the level of task complexity is gradually increased in accordance with the progress of learning so that minimum task performance requirements can be met throughout the learning process and that the learning process may not diverge as the level of task complexity increases.

Progressive learning is a dynamic process, since task assignments vary dynamically during the learning process. As shown in Figure 2, the system consists of a learner, a plant, a performance evaluator, and a learning organizer. The learning organizer determines the task complexity level appropriate for the learner on the basis of the task performance evaluation and provides the learner with a task program, referred to as the "curriculum". In consequence, there are two feedback loops involved in the progressive learning system: the curriculum assignment loop and the standard learning control loop. In designing this progressive learning system, we need to elaborate both the curriculum and the learning rule in an integrated and cohesive manner. There is a significant synergism between the two that may improve stability and convergence speed along with practical aspects, as will be explored later.

Central to this progressive learning is the design of the curriculum along with the learning control design. How to vary the task complexity level is a critical issue in designing a progressive learning system. In the following sections, we will address a particular way of varying the task complexity level. Namely, we vary a motion speed command, starting with a slow speed and ending in a high speed. The idea is to excite parameters of the system associated with quasi-static motion by driving the system slowly at the beginning. In this stage, the system learns only the quasi-static parameters, leaving other parameters unlearned. After the quasi-static parameters have been learned, motion speed is slightly increased. Then, the learning procedure is repeated for the higher speed and the system is relearned and retuned to the higher speed. As the motion speed increases, the dynamics becomes prominent and the parameters associated with dynamic motion are more excited. As long as the motion speed is increased gradually, the system does not have to learn all the parameters at the same time, but simply needs to refine the previous results within a limited range in the whole parameter space, which is often large. The learning parameter space is therefore excited gradually and learning proceeds progressively. This simplifies the learning problem significantly.

3. HIGH SPEED INSERTION TASK

3.1. *Task Description*

In this section, the concept of progressive learning is reduced to a concrete algorithm for high-speed robotic assembly. As shown in Figure 3-(a), the task is simply to insert a ball into a chamfered hole in an $x - y$ plane. The controller is given a nominal trajectory $\boldsymbol{x}_d(t) = (x_d(t), y_d(t))^T$. However, due to the uncertainty inherent in the assembly process, the hole is not precisely aligned with the trajectory and the ball often collides with a chamfer surface. Compliance control is necessary to cope with the geometric uncertainty of the assembly process, but is not sufficient for high speed insertion. For example, when the ball approaches a chamfer at high speed and collides with the surface, the quasi-static controller may not be able to prevent the ball from bouncing on the chamfer surface, which may lead to a failure of insertion. To avoid bouncing as well as to guide the ball correctly despite high speed, we need a dynamic control law such as full impedance control including damping and inertial terms along with the compliance or stiffness term. Such dynamic control laws contain a number of parameters to be tuned to a specific task process. It is a difficult job to find the optimal values in a large parameter space, particularly when all the parameters must be learned on-line in real time. It should be noted that a failure in high speed assembly may incur serious damage to the robot as well as to the parts and the environment. Even for the purpose of learning, fatal mistakes must be avoided at all times. Therefore, we intend to apply progressive learning to cope with these difficulties.

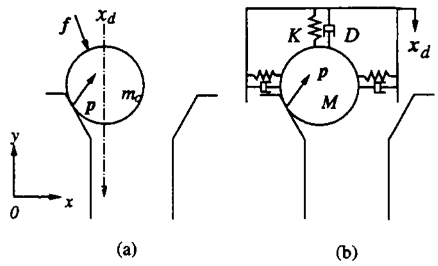


Fig. 3. Schematic diagram of impedance control

As shown in Figure 3-(b), a ball is held with an appropriate impedance. We begin by formulating the impedance control law in accordance with [4]. The motion of the ball of mass m_0 is governed by the equation of motion given by

$$\boldsymbol{f} + \boldsymbol{p} = m_0 \ddot{\boldsymbol{x}} \tag{1}$$

where $\boldsymbol{x} = (x, y)^T$ is the position of the ball with an inertial reference, $\boldsymbol{p} = (p_x, p_y)^T$ is the contact force acting on the ball, and $\boldsymbol{f} = (f_x, f_y)^T$ is the actuator's force to be controlled. The objective of impedance control is to emulate a desired mechanical impedance by controlling actuator force $\boldsymbol{f}$. The desired dynamics of the system shown in Figure 3-(b) is given by

$$\boldsymbol{p} = M\ddot{\boldsymbol{x}} + D(\dot{\boldsymbol{x}} - \dot{\boldsymbol{x}}_d) + K(\boldsymbol{x} - \boldsymbol{x}_d) \tag{2}$$

where $\boldsymbol{x}_d = (x_d, y_d)^T$ is the nominal trajectory, and M, D and K are the desired inertia, damping and stiffness matrices respectively. The external force $\boldsymbol{p}$ is measured by a force sensor attached to the end effector. From eqs.(1) and (2), we can derive the impedance control law given by

$$\boldsymbol{f} = m_0 M^{-1} D(\dot{\boldsymbol{x}}_d - \dot{\boldsymbol{x}}) + m_0 M^{-1} K(\boldsymbol{x}_d - \boldsymbol{x}) + (m_0 M^{-1} - I)\boldsymbol{p}$$

To formulate a learning algorithm, we need a means for evaluating control performance. In accordance with [11], we will define a performance index function, referred to as a reinforcement function. In robotic assembly, the objective of control is to mate a part with an uncertain fixture while minimizing the interference and conflict between the part and the fixture. In the ball insertion task, the robot is required to insert a ball into a hole with a minimum reaction force from the walls. At the same time, the controller must follow the nominal trajectory at least until the ball makes contact with a chamfer or a wall since the nominal hole position is the best initially available conjecture for the real position. Namely, the controller is required to follow the nominal trajectory closely while producing the smallest possible reaction force. Based on these considerations, we define the reinforcement function r, a performance index for the controller, as follows.

$$r = -[\rho_1\|\boldsymbol{x}_d - \boldsymbol{x}\|^2 + \rho_2\|\dot{\boldsymbol{x}}_d - \dot{\boldsymbol{x}}\|^2 + \rho_3\|\boldsymbol{p}\|^2] \quad (4)$$

where ρ_1, ρ_2 and ρ_3 are weighting factors of the individual terms. The problem is to learn the impedance parameters, K, D and M, in eq.(3) so that the above reinforcement can be maximized. Our approach to this learning problem is to increase the reference motion speed, $\dot{\boldsymbol{x}}_d$, progressively and repeat a learning procedure for different motion speeds so that the impedance matrices can be learned in sequence.

3.2. *The Progressive Learning Approach: Accommodating the Task Complexity Level by Motion Speed Scheduling*

The objective of motion speed scheduling is twofold: one is to prevent the learning process from diverging and the other is to maintain minimum task performance and avoid fatal mistakes and damage. As will be shown later, when a robot attempts to learn all the parameters simultaneously, the process tends to diverge or, even if it converges, searching for the optimal parameters can be a lengthy process. Furthermore, unless appropriate initial parameters are assigned, the control system may become unstable and even dangerous at high speed operations, and this may result in serious damage to the system. Therefore, learning must be initiated at low speed and the motion speed must be increased gradually as the robot gains control knowledge.

A. Slow Speed Motion

In progressive learning, we start with a slow speed mainly to excite the quasi-static terms in the impedance control law given in eq.(3). A gradient-following learning algorithm is applied to the reinforcement function given in eq.(4) to learn the control parameters. In this slow speed learning, only stiffness or compliance terms can gain the most information from the learning process. Since the damping and inertia terms are not fully excited at this slow speed, meaningful values cannot be acquired for those terms.

To execute this learning, initial values for impedance parameters K, D and M must be provided. These, however, need not to be accurate; one can use some positive matrices so that the robot can track the nominal trajectory $\boldsymbol{x}_d$ stably to perform a task at low speed. As learning proceeds, the stiffness parameter K will be updated toward the optimal stiffness, which maximizes the reinforcement function while changes in D and M remain small.

B. Medium Speed Motion

After the learning curve of stiffness K has reached a plateau or the reinforcement function has reached a certain threshold level, we increase the motion speed $\dot{\boldsymbol{x}}_d$ to a medium speed. As a result, the assembly process becomes more dynamic and the damping term D in eq.(3) now becomes highly excited. We can use the same learning algorithm as in low speed learning to learn the new parameters. The optimal stiffness that was learned at the slow speed learning must be used as the initial values for K in this phase of learning. The damping and inertial terms in the previous phase are also used as the initial values for D and M. In the beginning of learning, the increased speed temporarily deteriorates the performance of the controller, but the controller adapts itself to the higher speed operation by quickly learning the damping matrix D. Important to note here is that we can make a smooth transition from the slow motion operation to the medium speed operation by succeeding the previously learned parameters as the initial values for the next learning procedure. In other words, if we were to choose initial values randomly, the controller's performance would be extremely poor at the beginning and, as a result, would cause some damage to the task environment. With progressive increase of motion speed, we can avoid these problems and find the optimal controller faster and more effectively.

C. High Speed Motion

After learning the optimal damping has been completed, we further increase the motion speed to the highest. This increase of motion speed excites the system in a broader range, and makes the dynamic characteristics more prominent. The inertia term becomes dominant in this phase. As at the medium speed, the impedance matrices acquired in the previous phase are used for the initial impedance matrices. By doing so, we can avoid the instability and damage to the system that could otherwise be caused. After temporary deterioration of performance caused by the increase of the motion speed, the controller will smoothly reach the optimal impedance matrices for the high speed motion.

4. IMPLEMENTATION AND SIMULAITON

In this section, we implement the above progressive learning of impedance parameters and conduct simulation experiments to demonstrate the effectiveness of the proposed method. Various simulation results are then provided to verify the arguments given above.

4.1. *Implementation*

In this paper, we implement the progressive learning method by using the Adaptive Reinforcement Learning Algorithm(ARL) [11]. The ARL algorithm applies a perturbation/correlation technique to learning an internal model. The estimated gradient and parameter changes based on the internal model are less erratic and more robust against uncertainties and noise compared with the general model-based learning algorithms. The details of the ARL algorithm are provided in [11]. In the ARL algorithm, a radial-basis function network is used to represent an internal model. As discussed in the previous section, locally tunable networks "synergize" progressive learning, since the system is excited locally and gradually. Even if the accuracy of the model network is limited at the beginning, the accuracy can be quickly improved during on-line learning due to scheduled excitation.

In this simulation experiment, we use a spring-damper model to simulate the chamfer surface and compute the impact of collision between the ball and the chamfer surface. With this model, we can calculate the contact force or impact force p_n in the direction normal to the chamfer surfaces. Friction p_t on the chamfer surfaces is given by $p_t = \mu p_n$, where μ is a friction coefficient.

In each learning iteration, the controller is given a nominal trajectory x_d. The trajectory is parallel to but deviated randomly from the center line of the hole so that the ball always collides with the chamfer surface. The impedance parameters involved in the controller are tuned in real time using the Adaptive Reinforcement Learning Algorithm. In this simulation experiment, we conducted 300 iterations of learning divided into three phases: the first 100 iterations in a slow speed, the second 100 iterations in a medium speed and the last 100 iterations in a high speed.

4.2. *Simulation Results*

Figure 4 shows the transitions of the impedance parameters as well as the reinforcement over the whole learning iterations. Note that for each iteration we averaged the reinforcement over the time period of the iteration. Note also that the erratic behavior of the learning curves is caused by the randomness involved in each iteration as well as in the Adaptive Reinforcement Learning Algorithm. The details of the learning procedure follow.

As described in the previous section, learning started with a slow motion speed to excite only the quasi-static terms of the system. The velocity command $\dot{x}_d$ was set to $(0, -0.3\ \mathrm{m/s})^T$. The

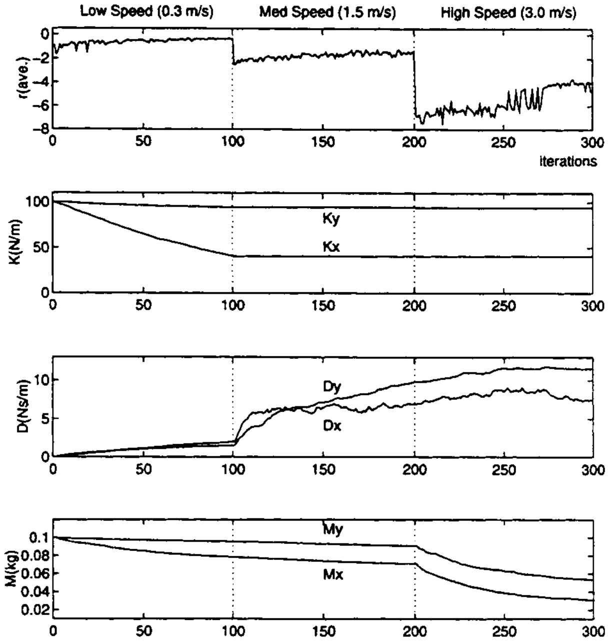


Fig. 4. Learning histories of reinforcement r and stiffness K, damping D and inertia matrices M over 300 iterations

initial stiffness matrix was given by:

$$K_0 = \begin{bmatrix} 100\ \mathrm{N/m} & 0 \\ 0 & 100\ \mathrm{N/m} \end{bmatrix} \tag{5}$$

These stiffness parameters, or position gains, are large enough to follow the desired trajectory. The damping matrix D was initialized with small positive values and the inertia matrix was initialized as the plant inertia matrix $m_0 I$.

One hundred iterations of learning were performed in this phase. As shown in Figure 4, the stiffness in the x direction decreased remarkably to $47.41 N/m$ while that in the y direction decreased just slightly to $96.3 N/m$ during the 100 iterations. This shows that the robot has to hold the ball compliantly in the x direction to comply to the chamfer surface and stiffly in the y direction to follow the nominal trajectory toward the bottom of the hole. This result is compatible with the argument that was derived by [10] and other researchers. On the other hand, the damping and inertia parameters are almost unchanged, showing that these terms are not excited during this slow speed operation. From this observation, we can draw the conclusion that in the slow speed operation the damping and inertia parameters cannot gain useful information. The average of the reinforcement values converged to -0.5.

Secondly, we increased the motion speed to $\dot{x}_d = (0, -1.5\ \mathrm{m/s})^T$ to excite the system more dynamically, and conducted another 100 learning iterations. As shown in Figure 4, in the beginning of this learning phase, the reinforcement value decreased discontinuously from -0.5 to -2.6 due to the increase of speed. Namely, the controller trained in the slow speed operation could not perform satisfactorily in this faster motion. However, as learning proceeded, the performance was quickly improved and the reinforcement again started to increase. In this phase, the damping parameters increased steadily and significantly while the stiffness parameters remained almost constant over the iterations. It appears that the stiffness parameters had already been learned in the first learning phase and the controller did not gain any additional information at the faster motion speed. The inertia matrix M was again almost unchanged in this learning phase.

Finally, we increased the motion speed to the maximum $\dot{x}_d = (0, -3.0 \text{ m/s})^T$. When the task was first performed at maximum speed, there was a large amount of bouncing on the chamfer surface resulting in huge impact forces. In order to reduce the bouncing and the impact forces, the inertia term must be significantly modified. As shown in Figure 4, at this motion speed, the inertia terms were very vigorously excited and the corresponding parameters moved most significantly and converged to the final values. The damping parameters also varied in this learning phase while the stiffness parameters again remained unchanged. Note that after the learning the inertia in the x direction became much smaller than that in the y direction, which is also compatible with the theoretical conclusion derived in [1]

As shown above, by increasing the motion speed progressively, the controller can learn a better impedance smoothly and stably while maintaining a minimum performance level. To compare the progressive learning method with its traditional counterpart, we conducted another set of learning simulations. In the traditional non-progressive learning method, a constant motion speed, 3.0 m/s, was used throughout the learning operations. Figure 5 shows the learning curves of the two methods. As shown in the figure, the control performance i.e. the reinforcement value, is erratic and significantly lower than that of the progressive learning method. Important to note is that the progressive learning method allows the controller to maintain a certain performance level throughout all the learning iterations while non-progressive learning sometimes exhibits intolerable, poor performance, which may incur significant damage to the system.

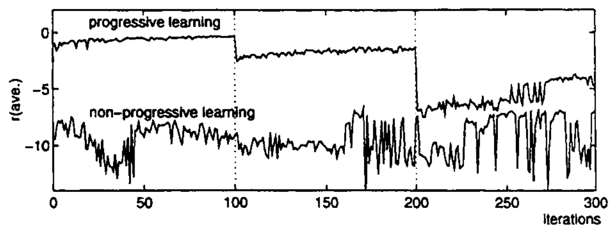


Fig. 5. Comparison between progressive and non-progressive learning methods

5. CONCLUSION

In this paper we have proposed the concept of progressive learning and a scheduled excitation method by varying a motion speed command. Based on this method, we developed an impedance learning algorithm for high speed insertion. We also fully manifested the mechanisms of progressive learning and examined why the progressive learning method works successfully for impedance learning. The progressive impedance learning method was implemented by using the adaptive reinforcement learning algorithm, and simulation experiments were conducted to show the effectiveness of the progressive learning method.

6. REFERENCES

[1]H. Asada and Y. Kakumoto, "The Dynamic Analysis and Design of a High-Speed Insertion Hand Using the Generalized Centroid and Virtual Mass," ASME Journal of DSMC, vol. 112, pp. 646-652, 1990

[2]E. W. Bai and S. S. Sastry, "Global Stability Proofs for Continuous-Time Indirect Adaptive Control Schemes," IEEE Trans. of Auto. Cont., 1987

[3]S. Boyd and S. S. Sastry, "Necessary and Sufficient Conditions for Parameter Convergence in Adaptive Control," Automatica, vol. 22, pp. 629-639, 1986

[4]N. Hogan, "Impedance Control: An Approach to Manipulation: Part I-III," ASME J. of DSMC, vol. 107(1), 1985

[5]R. Horowitz, et al., "Convergence Properties of Learning Controllers for Robot Manipulators," Proc. of the Japan-USA Symp. on Flex. Auto., Japan, July, 1990

[6]M. I. Jordan and D. E. Rumelhart, "Forward Models: Supervised Learning with a Distal Teacher", Cognitive Science, 16, 1992

[7]K. S. Narendra and A. M. Annaswamy, "Persistent Excitation of Adaptive Systems," Int. Journal of Control, vol. 45, pp. 127-160, 1987

[8]K. S. Narendra and A. M. Annaswamy, Stable Adaptive Systems, Prince Hall, 1989

[9]T. Poggio and F. Girosi, "A Theory of Networks for Approximation and Learning," A.I. Memo No. 1140, AI Lab., MIT, 1989

[10]D. E. Whitney, "Force Feedback Control of Manipulator Fine Motions," ASME J. of DSMC, vol. 99(2), 1977

[11]B.-H. Yang and H. Asada, "Adaptive Reinforcement Learning and Its Application to Robot Control", DSC-Vol. 49, ASME WAM, 1993

THE DESIGN OF PERCEPTUALLY AUGMENTED DISPLAYS TO SUPPORT INTERACTION WITH DYNAMIC SYSTEMS

Alex Kirlik

Center for Human-Machine Systems Research
School of Industrial & Systems Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332 USA
kirlik@chmsr.gatech.edu

Abstract: Perceptually rich information displays, depicting concrete system representations, are increasingly used to support operators of complex human-machine systems. Although much of the psychological literature provides an optimistic picture of the effects of these representations, the relevance of this literature to the design of complex system interfaces is less direct than it would appear. We identify three distinctions that should be considered when evaluating the applicability of the display design literature to actual design problems. Using these distinctions, we review the literature and two of our own experiments with an eye toward identifying the cost-benefit tradeoffs associated with a shift from abstract to more concrete displays, and discuss a number of caveats that should be kept in mind when providing operators with perceptually enriched display representations. We identify needs for future research necessary to increase the relevance of scientific research to actual interface design problems.

Keywords: Displays, Human-Machine Interface, Perception, Decision Making, Skill

1. INTRODUCTION

A clear trend in human-machine interface design is a move toward increasingly concrete, perceptually-rich, information displays. Empirical research on the effects of enriched displays, however, provides mixed and in some cases inconsistent results concerning the benefits and costs of these representations. We review the literature and discuss two of our own recent experiments on the effects of perceptually enriched displays in order to resolve these discrepencies, and to provide a more coherent picture of the effects of increasing the concreteness of the system representation provided to the human operator.

Our analysis of the empirical evidence relies on a set of three distinctions that should be considered when evaluating the relevance of display research findings to a particular design problem or performance issue. A failure to make these distinctions has, in our opinion, contributed to the appearance of inconsistency and incoherence in the literature on both display effects on performance and prescriptive guidelines for interface design.

The first distinction we discuss concerns the role of the displayed representation: the representation can be either the object of interaction itself, or it can be merely the window to a system or environment which is the actual object of interaction. The majority of empirical research on display effects falls under the first category, while most interface design problems also raise questions associated with the actual *representational* role of a display. The second distinction we discuss concerns the type of performance measures used to evaluate display quality: performance measures can assess solely the quality of human interaction with the displayed representation, or they can assess the quality of the operator's ability to perform when the display is not available. Again, most research primarily uses the first category of performance measures, while actual design problems also raise questions concerned with what the operator learns about the domain *beyond* the use of information explicitly represented at the interface. Finally, we discuss a distinction concerning the type of system or domain model that serves as the basis for the displayed representation: for some domains or systems, complete and certain models are available, while for others, the best available system model is either incomplete, uncertain, or both. As is the case with the previous distinctions concerning the role of the display and the relevant measures of performance, there is also often a considerable difference between the domain models employed in empirical research and the models available to designers of actual human-machine systems.

In the following, we first review the literature on the effects of perceptually enriched displays, and provide an overview of our own research in this area. We then consider the empirical data from the standpoint of the three distinctions outlined above, with an eye toward resolving some of the apparent inconsistencies in the empirical findings, and perhaps more importantly, also resolving some of the inconsistencies between the prescriptive design principles that have resulted from this research.

2. CONCRETE REPRESENTATIONS: COSTS & BENEFITS

Benefits of graphical or pictorial representations on human information processing have been documented in a variety of tasks: for example, geometry theorem proving (Koedinger and Anderson, 1990), physics (Larkin and Simon, 1987), puzzle solving (Zhang and Norman, 1994), operations research (Moody and Kirlik, in press), understanding device behavior (Hegarty and Simms, in press), economics (Tabachneck, Leonardo, and Simon, 1994), and fault diagnosis (Kieras, 1992; Vicente and Rasmussen, 1990). These researchers generally agree that concrete, graphical representations create information processing economies associated with information search and perceptual inference.

There is less agreement on a theoretical explanation of the benefits of graphical displays. The explanations that have been proposed range from efficiencies describable via production system models (e.g., Larkin and Simon, 1987; Koedinger and Anderson, 1990; Casner, 1991), the facilitation of mental model acquisition and use (e.g., Bauer and Johnson-Laird, 1993; Hegarty and Simms, in press; Kieras, 1992), to more radical explanations that may require significant extensions to the traditional information processing framework. For example, Zhang and Norman (1994) argue that understanding the benefit of external problem representations will require a theory of cognition as "distributed" across the human-environment system, and Vicente and Rasmussen (1990) argue that a central benefit of graphical displays is in enabling the direct perception of affordances

Much of the research cited above has resulted in findings that are consistent with many peoples' intuitions that more concrete visual representations are superior to more abstract representations, that "a picture is worth a thousand words." Perhaps even more interesting, then, is research that has documented situations or tasks in which increasing the concreteness of a representation results in decreased human performance.

For example, experiments by Kieras (1992) provide evidence suggesting that the visual complexity caused by a high-fidelity display of system structure can sometimes trade off with the benefits of a graphical representation. In addition, Coury and Pietras (1989) describe an experiment in which a graphical display resulted in performance that was inferior to a digital display in a process control task. Gersh and Hamill (1992) describe the results of an experiment on human supervisory control performance in which graphically displaying a decision variable merely caused operators to spend more time and effort manipulating the variable, but with no resulting performance improvement. Finally, on the basis of their own experiments, Larkin and Simon (1987) put forth the counter-intuitive notion that although graphical displays may enhance information search and recognition, they have no effects on the efficiency or accuracy of inference itself.

Wickens (1992) has raised the concern that while increasingly concrete displays may improve a performer's ability to interact with a system, such an approach may actually decrease the performer's ability to learn deeper knowledge of the domain that could be used when the display was made unavailable. Wickens reviewed a number of empirical studies in support of his position. Each of those studies has demonstrated either better retention of information, or better performance, for a more abstract rather than concrete display after the removal of augmented information. One could summarize this series of empirical findings in terms of a "no pain - no gain" theory of skill acquisition or learning: although a task might be made easier by providing a performer with environmental support, providing less support will require the performer to learn deeper or more abstract knowledge about the task domain that the supported performer would never learn.

Some researchers have raised a closely related concern about the potential costs of increasingly concrete representations in human-machine systems. Specifically, it may be the case that a perceptually rich interface design might have the unintended side effect of focusing the operators' attention toward interaction with the interface, and away from the operator's true task. In most complex human-machine systems, the operator's actual task is not to control an interface, but rather, to effectively control a system which is merely *represented* by the interface. In the context of the design of process control displays, for example, Montmollin has suggested:

> If "sophisticated" engineers try to "facilitate" (in their opinion) the work, by transfer to the screen(s) the totality of information, which moreover is modified and interpreted by some intermediate software, it could happen that for the operator the task is no longer the control of the process, but more the interpretation and control of the interface, which becomes the "object" of the operator's work. This is not always the best solution! (Montmollin, 1991, p. 100).

One seldom considered, yet possible, benefit of abstract information displays may be the clarity with which it can be seen that the displayed representation is not the controlled system itself; that the "map" is not the "territory." As Montmollin observes, in our rush to exploit available technology to achieve some of the documented benefits of graphical displays, it may be that a price will be paid in terms of the operator's ability to reason beyond the available representation.

Relating to this issue, research by Fischoff, Slovic, and Lichtenstein (1978) has demonstrated that subjects given a graphical (tree diagram) representation of a decision problem fail to consider contingencies not explicitly included in the representation. In addition, research by Smith (1990) has demonstrated that subjects presented with a correct, yet incomplete, graphical representation of a decision making task performed less well than subjects who were presented with no graphical representation at all. While both subject groups in Smith's experiment received a verbal description of the decision problem, supplementing the verbal representation with a more concrete, graphical representation actually decreased decision making accuracy. These results raise the concern that any incompleteness in a task representation may be even more problematic in designing graphical displays, than in designing traditional, more abstract displays.

Finally, a somewhat related concern about the use of increasingly concrete task representations was expressed by Kirlik, Miller, and Jagacinski (1993) in the context of human supervisory control. Quoting directly:

> .... we must point out that very little is known about the tradeoffs that occur when a cognitively intensive task is changed into a predominantly perceptual task through the use of information-rich task representations. For example, cognitive "biases" may be traded for the limitations of perception, and we may reduce the frequency of errors resulting from highly flexible cognitive activities (i.e., "mistakes") only to find we have increased the potential for errors resulting from less flexible perception-action mechanisms (i.e., slips). It is certainly premature to advocate a comprehensive interface design framework until more is known about such tradeoffs (p. 950)

Research on human interaction with perceptually rich, graphical displays has yet to result in a good understanding of such tradeoffs. Below, we give a brief overview of two experiments we have recently performed in order to investigate these issues.

3. TWO EXPERIMENTS ON AUGMENTED GRAPHICAL DISPLAY SUPPORT

We have used a microworld methodology to investigate the effects of enriched perceptual displays on human skill acquisition and performance in two dynamic decision making and control tasks. We first describe the analysis and design approach underlying both of these experiments.

3.1 Ecological Task Analysis

We have advanced an approach to the design of perceptually augmented displays called the Ecological Task Analysis (ETA) framework (Kirlik, in press). ETA is proposed as a solution to the problem of representation design (i.e., identifying the specific information to be displayed) for supporting skilled human interaction with dynamic systems. Unlike related approaches such as Ecological Interface Design (Vicente and Rasmussen, 1990), ETA does not require detailed models of the dynamics of a controlled system to serve as the display representation. Instead, the focus within ETA is on the human operator's repertoire for action, and the environmental or system constraints that specify when each action within the repertoire is appropriate given the environmental state. The main assumption underlying ETA is that display design should focus on creating forms to communicate the constraints underlying the behavior of the task environment that are directly relevant to the selection of action.

ETA's focus on displaying the environmental constraints on productive action was motivated by studies of skilled interaction with complex dynamic systems. These studies suggest that operators do not select control actions based upon reasoning with a detailed model of the controlled system (Kirlik, Miller, and Jagacinski, 1993; Brehmer, 1990; Brehmer and Allard, 1990). Rather, human strategies in such tasks appear to be highly reliant upon perceptual or pattern recognitional activities that are not explicitly model-based (also see Rouse, 1983; Reason, 1988). This relatively efficient, heuristic mode of control, however, can only be productive if perceptual information is available to specify all the relevant constraints upon action selection. ETA provides guidance to ensure that information regarding all relevant environmental constraints are included in a display representation. In short, the goal of ETA is to ensure a display provides information indicating what actions are and are not productive, given the dynamic environmental state; i.e., to allow the operator to be able see what he or she can do.

3.1 Experiment 1: ETA and StarCruiser

StarCruiser is a dynamic decision and control task in the form of a space-related videogame. The operator pilots a spaceship through a galaxy populated by a number of solar systems and associated planets, all shown on a graphical computer display. The operator's goal is to deploy a variety of tools to the planets to collect valuable resources, and to return these resources to a starbase, at which time points are earned for the resources collected.

The experiment involved using ETA to motivate the design of various forms of perceptual augmentation to support operator performance in StarCruiser. The experiment compared the performance of subjects using either a baseline or perceptually augmented display over 18 sessions (days). The augmented display was identical to the baseline display, except that it included additional graphical information that more directly specified the constraints in the task environment that determined what actions were feasible at each point in time. A detailed description of this experiment appears in Kirlik, Kossack, and Shively (1994). Below, we briefly summarize the key findings of this experiment.

Key Findings. Perceptual augmentation in StarCruiser accelerated skill acquisition but did not affect peak performance in the late sessions. However, a simultanteous cognitive task (mental arithmetic) was administered in the penultimate sessions and this manipulation appeared to impair StarCruiser performance for the baseline display group, while it did not impair StarCruiser performance for the group using the augmented display. Finally, in the last sessions the augmented group subjects performed using the baseline display, to assess the effects of the augmented perceptual information becoming unavailable. Lack of augmented information did not impair performance.

3.2 Experiment 2: ETA and EJSTARS

The second experiment was conducted using a laboratory task known as the Extended Joint Surveillance and Attack System (EJSTARS). A detailed description of this task and experiment is presented in Krosnick (1994). EJSTARS is a hypothetical, yet technologically feasible, integration of unmanned aerial vehicles (UAVs) in the current U.S. Air Force/Army Joint Surveillance and Attack System (JSTARS). The operational JSTARS system combines ground surveillance radar with sophisticated command and control systems aboard a militarized 707 (E-8) aircraft. Our laboratory simulation was designed to capture the decision making characteristics of the operational system, as well as to maintain face validity with that system. EJSTARS presented laboratory participants with a task complexity on par with many popular computer-based, or arcade video games. The EJSTARS operator's main task is to identify initially unknown vehicles moving on a map display, and to take defensive action upon only those vehicles posing a certain threat to the operator's own resources.

The experiment involved using ETA to motivate the design of various forms of perceptual augmentation to support operator performance in EJSTARS. The experiment compared the performance of subjects using either a perceptually augmented or baseline display over 18 sessions (days). Each of the four specific forms of perceptual augmentation included on the augmented EJSTARS display was associated with one of the four decision making constraints indicating whether or not the operator should engage a particular enemy vehicle. These constraints were: 1) Penetration: the operator should not engage a vehicle whose weapons were not sufficiently powerful to penetrate the armor of the operator's own (friendly) resources; 2) Range: the operator should not engage a vehicle that cannot move within its weapons range of a friendly resource; 3) Locomotion: whether or not a vehicle can move within its weapons range of a friendly resource is determined by terrain properties such as steepness, iciness, etc.; and 4) Priority: in the event of multiple, simultanous threats, the operator should engage the vehicle threatening the highest priority friendly resource. The perceptually augmented display included graphical forms to allow these constraints to be readily perceived from the graphical map display, while the baseline display required subjects to combine alphanumerically displayed information with the map information in order to assess these constraints.

Key Findings: Perceptual augmentation in EJSTARS accelerated early skill acquisition but did not affect peak performance in the middle and late sessions. However, when task demands were increased in two late sessions (the number of vehicles in the environment was doubled), the performance of the baseline display subjects was more impaired than was the performance of the augmented display subjects. In the last sessions, the augmented display group subjects performed with the baseline display, to assess the effects of the augmented perceptual information becoming unavailable. Lack of augmented information did not impair performance.

These findings echo the results of the previous experiment in the StarCruiser task. The decision making demands in EJSTARS were designed to be much more complex than those in StarCruiser, because we hypothesized that the lack of a main effect of augmentation in StarCruiser on peak performance could be due to a ceiling effect, and we thus wanted to see if that finding would replicate in a more demanding task environment. We were surprised that this finding did replicate, and we thus performed additional analyses to understand why the augmented display did not improve peak performance in addition to the rate of skill acquisition. In short, these analyses indicated that the null effect on peak performance was due to a combination of both positive and negative effects of the various forms of enhanced information used on the augmented display. Augmentation associated with deterministic task constraints increased decision accuracy, while augmentation associated with probabilistic decision constraints decreased decision accuracy. Subjects using the augmented display overweighted the certainty (specifically, the predictive value) of the graphical information associated with probabilistic task constraints, while subjects using the baseline display better weighed the limited predictive value of the same information when it was displayed in alphanumerical form. In summary, we suspect that the concreteness with which probabilistic information was presented on the augmented display prompted subjects to treat this information as if it had much greater predictive value than it actually had.

4. DESIGN IMPLICATIONS

What are the implications of all these empirical findings for the design of interfaces for human-machine systems? One approach toward answering this question would be to integrate these findings into a set of prescriptive guidelines or some optimal design framework given current knowledge. However, our experience suggests that such an integration would be of limited value. First, despite the apparently impressive number of studies on the issue of concrete versus abstract representations, the design space is so large that the current findings represent only a very sparse sampling of this space: a universal design framework is probably premature given current knowledge. In addition, it should be obvious that the rapid proliferation of increasingly concrete representations, not only in complex systems, but also in areas such as education and training, is hardly due to educating display designers on some shared psychological framework. Instead, the proliferation of these interfaces is mainly due to the availability of technology which increasingly allows individual designers to create intuitively appealing (typically more concrete) display designs.

Given these observations, we believe it may be more valuable to pursue the path suggested by Heisenberg, who noted that "an expert is someone who knows some of the worst mistakes that can be made in his subject, and how to avoid them." In conclusion, then, we discuss caveats that should be kept in mind when considering applications of enriched displays, and identify some human performance problems we expect to see more often as a result of the proliferation of these interfaces.

4.1 Treating the Map as the Territory

A troubling trend in the literature reviewed above is that increasing the concreteness of a visual represenation may promote a tendency for humans to treat the displayed representation as if it, and not the represented system or environment, is the actual target of interaction. Although an impressive array of cognitive science research findings document benefits of graphical displays, little of this research provides subjects with both a problem representation *and* a represented problem situation (e.g., a diagram of a pulley-weight system, *and* an actual pulley-weight system that must itself be manipulated to solve a particular problem). The quotation from Montmollin cited previously speaks specifically to this concern. More research is certainly needed to understand the optimal design of representations that allow the operator to reason effectively with the representation, and also ensure that the operator can and will be able to view the display as "only" a model when it becomes crucial to do so.

4.2 Incomplete/Uncertain Representations

A bias to treat the map as the territory may be most costly when the displayed representation is either incomplete, uncertain, or both. Again, cognitive science research mainly uses formal systems (e.g., geometry, puzzles) as laboratory tasks, for which complete, determinstic representations can be created. Rarely, however, do designers of complex system interfaces have available a complete, certain model of the controlled system to serve as a display representation. Some of the previously reviewed literature suggests that subjects presented with more concrete representations are less able to reason beyond those representations (to consider unrepresented decision options or states of nature) than are subjects given more abstract representations. Our own findings in the EJSTARS experiment indicated a benefit for perceptual augmentation associated with deterministic task criteria, but augmentation associated with probabilistic criteria decreased performance. More research is needed to understand the cost/benefit tradeoffs that occur when providing an increasingly concrete representation based on an incomplete or probabilistic system model.

4.3 Withdrawn Representations

Is a representation that fosters optimal performance in the presence of the display the same representation that also fosters optimal performance when the performer must work without the display? The previously cited paper by Wickens is perhaps the most cogent discussion of this tradeoff, in the context of displays for educational systems. In education, it is natural to view the role of a learning environment to promote learning that transcends the learning environment itself. Similar concerns are also relevant to the training and operation of complex human-machine systems. Our own research has yielded promising findings on how to exploit perceptual augmentation to accelerate skill acquisition for training systems, without the subsequent removal of augmenation hindering performance in the unenhanced environment.

5. CONCLUSION

Emprical research has only sparsely sampled the design space with respect to the issue of concrete versus abstract display representations. Features of the design space of actual system interfaces (e.g., incomplete, uncertain system models) have actually been undersampled in scientific research, and thus a naive reading of the literature may result in an overly-optimistic prediction of the effects of increasingly concrete displays. We discussed a number of distinctions that must be made to evaluate the relevance of laboratory research for design problems. Using these distinctions, we noted some caveats that should be considered in the design of enriched displays, and outlined needs for further research to increase the applicability of research findings to system design.

REFERENCES

Bauer, M. I. and Johnson-Laird, P.N. (1993). How diagrams can improve reasoning. Psychological Science, 4, 372-378.

Brehmer, B. and Allard, R. (1990). Dynamic decision making: The effects of task complexity and feedback delay. In J. Rasmussen, B. Brehmer, and J. Leplat (Eds.), Distributed Decision Making: Cognitive Models for Cooperative Work (pp. 319-334). New York: John Wiley & Sons, Ltd.

Brehmer, B. (1990). Strategies in real-time decision-making. In R. Hogarth (Ed.), Insights in Decision Making (pp. 262-279). Chicago, IL: University of Chicago Press.

Casner, S.M. (1991). Task-analytic approach to the automated design of graphic presentations. ACM Transactions on Graphics, Vol. 10, No. 2, 111-151.

Coury, B.G. and Pietras, C.M. (1989). Alphanumeric and graphic displays for dynamic process monitoring and control. Ergonomics, Vol. 32, No. 11, 1373-1389.

Fischoff, B., Slovic, P., and Lichtenstein, S. (1978). Fault trees: Sensitivity of estimated failure probabilities to problem representation. Journal of Experimental Psychology: Human Perception and Performance, 4, 330-334.

Gersh, J.R. and Hamill, B.W. (1992). Cognitive engineering of rule-based supervisory control systems: Effects of concurrent automation. *Proceedings of the 1992 IEEE Int'l Conference on Systems, Man, and Cybernetics.* Chicago, IL,

Hegarty, M., and Simms, V.K. (in press). Individual differences in mental animation during mechanical reasoning. Memory and Cognition.

Kieras, D. (1992). Diagrammatic displays for engineered systems: Effects on human performance in interacting with malfunctioning systems. International Journal of Man-Machine Studies, 36, 861-895.

Kirlik, A. (1994). Requirements for psychological models to support design: Toward ecological task analysis. In J.M. Flach, P.A. Hancock, J.K. Caird, and K.J. Vicente (Eds.), An Ecological Approach to Human Machine Systems I: A Global Perspective. Hillsdale, NJ: Erlbaum.

Kirlik, A., Kossack, M. F., and Shively, R.J. (1994). Perceptual augmentation to support skill acquisition in a dynamic decision and control task. Manuscript submitted for publication.

Kirlik, A., Miller, R.A., and Jagacinski. R.J. (1993). Supervisory control in a dynamic and uncertain environment: A process model of skilled human-environment interaction. IEEE Transactions on Systems, Man, and Cybernetics, Vol. 24, No.4.

Koedinger, K.R., and Anderson, J.R. (1990). Abstract planning and perceptual chunks: Elements of expertise in geometry. Cognitive Science, 14, 511-550.

Krosnick, B.W. (1994). Perceptual Augmentation to Support Skill Acquisition and Robust Decision-Making and Control Skills. Unpublished M.S. Thesis, School of Industrial & Systems Engineering, Georgia Institute of Technology, Atlanta, GA.

Larkin, J., and Simon, H.A. (1987). Why a diagram is (sometimes) worth ten thousand words. Cognitive Science, 11, 65-99.

Montmollin, M. (1991). Analysis and models of operators' activities in complex natural life environments. In J. Rasmussen and H.B. Anderson (Eds.), Research Directions in Cognitive Science: European Perspectives, pp. 95-112. Hillsdale, NJ: Erlbaum.

Moody, L.E. and Kirlik, A. (in press). Using visualization to support human problem solving in combinatorial optimization. IEEE Transactions on Systems, Man, and Cybernetics.

Reason, J. (1988). Modelling the basic error tendencies of human operators. Reliability Engineering and System Safety, 22, 137-153.

Rouse, W.B. (1983). Models of human problem solving: Detection, diagnosis, and compensation for system failures. Automatica, 19(6), 613-625.

Smith, G.F. (1990). Representational effects on the solving of an unstructured decision problem. IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-19, 1083-1090.

Tabachneck, H.J.M., Leonardo, A.M., and Simon, H.A. (1994). How does an expert use a graph? A model of visual and verbal inferencing in economics. Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society, pp. 842-847. Hillsdale, NJ: Erlbaum.

Vicente, K.J. and Rasmussen, J. (1990). The ecology of human-machine systems II: Mediating 'direct perception' in complex work domains. Ecological Psychology, Vol. 2, 207-250.

Wickens, C.D. (1992). Virtual reality and education. *Proceedings of the 1992 IEEE Int'l Conference on Systems, Man, and Cybernetics.* Chicago, IL, 842-847.

Zhang, J. and Norman, D.A. (1994). Representations in distributed cognitive tasks. Cognitive Science, 18, 87-122.

A CASE-BASED DESIGN BROWSER TO FACILITATE REUSE OF SOFTWARE ARTIFACTS

Jennifer J. Ockerman and Christine M. Mitchell

Center for Human-Machine Systems Research
School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332-0205

Abstract: Large corporations and government agencies often incur unnecessary costs and errors in conceptual software design due to the lack of an 'institutional memory' and the specialization of systems. A *design browser* is introduced as the foundation of an institutional memory to support software reuse. The Design Browser is a combination of an artificial intelligence paradigm known as case-based reasoning and a conceptual framework that presents developers a unified structure through which to understand, compare, and contrast related design experience. The case-based system contains design information about various existing software designs. The conceptual framework provides software developers with a common vocabulary and a frame of reference in which to view the existing software. The Design Browser and empirical results of its evaluation are presented.

Keywords: software specifications, design, system design

BACKGROUND

A full description of the Design Browser's functions can be found in (Ockerman et al., 1994; Ockerman, 1995). A summary of those functions is provided to support the presentation of the evaluation results.

1. INTRODUCTION

Many large corporations and government agencies are responsible for the design and development of large software systems. Often there are several generations of these systems or a general function that is performed for specific applications. Due to the similarity over generations or among specific applications, reuse suggests itself as a natural technique to facilitate increasingly efficient and cost effective systems. That is, a new system that reuses an existing system at some level (e.g., conceptual design, requirements, code, interface objects, or functionality) should be less expensive to build than a system that is designed from scratch (Prieto-Diaz, 1993).

Two potential reasons for the lack of reuse, particularly at the conceptual design level, are: (1) lack of 'institutional memory' about related system designs; and (2) the belief on the part of the design team that the current application is *unique* and thus, would not benefit from the experience of designers of similar systems.

Without an institutional memory, that is, a method of collecting and storing design experience that is accessible to subsequent designers, it is difficult, if not impossible, to reuse prior design experience and solutions (Parnas et al., 1989; Curtis, 1989; Chen and Lee, 1993). Moreover, the specialization of system software, i.e., new generations or related but not duplicate systems, emphasizes the *differences* between applications rather than the *similarities*. This leads to situations in which the design team holds the belief that each new version of a system *needs* to be designed from scratch (Lanergan and Grasso, 1984; Cusumano, 1991; Mili et al., 1994). The result is that each new system evolves in isolation, with its own concepts, vocabulary, and implementation. Application-specific features permeate the design process from conceptual design, to functional specification, to lines of source code. Even if a mechanism for 'institutional memory' was available, the application-specific concepts and language of each design project detract from the ability of subsequent design teams to *recognize* similarities and reuse relevant design experience.

2. A SOFTWARE DESIGN BROWSER

For the conceptual design stage, this research proposes a *Design Browser* to help designers access and use previous design experience. The Design Browser offers a method of finding, browsing (e.g., identifying, comparing, and contrasting), and learning from previous designs. The Design Browser provides a foundation of an 'institutional memory'--a computational structure in which to store experience from previous designs--through an artificial intelligence paradigm known as case-based reasoning. To facilitate the accumulation and rapid assimilation of design experience, the Design Browser includes a conceptual framework to organize features and functions that are common across a class of design applications.

Case-based reasoning combines a *cognitive model* describing how people use and reason from past experience and a *technology* for finding and presenting such experiences (Domeshek and Kolodner, 1992). The cognitive model is a memory-centered model; "the basic idea is that people are good at figuring out what to do in new situations largely because they are able to remember and adapt things they did (or saw others do) in similar previous situations (Domeshek and Kolodner, 1992, p. 203)."

The Design Browser is based on the assumption that access to examples of existing designs and experiences from related projects will facilitate reuse. Thus, its case base consists of stories and examples from past design experiences. The concept of a Design Browser further proposes that effective decision support also requires a unifying framework through which to view, extract, and integrate past experiences. The Design Browser's conceptual framework provides designers with a set of concepts and common vocabulary to specify the current design that emphasize its similarities to past designs.

2.1 Conceptual Framework

The goal of the Design Browser is to foster the view that instantiations of a class of systems are more similar than different. The Design Browser organizes information around features common to the set of design applications by means of its conceptual framework. The framework is communicated to the user via an intuitive, graphical interface. The Design Browser stores past design experience as cases; the interface and conceptual framework provide the user access to the case-based information.

The Design Browser's conceptual framework has two hierarchical decompositions: one a whole-part and the other functional. Software can be viewed in much the same way as physical artifacts. Software, like other artifacts, is composed of parts which operate together to achieve a goal. It has been shown that all artifacts can be simultaneously decomposed into whole-part and functional structure (Domeshek et al., 1993; Rasmussen, 1987; Neighbors, 1984). The simultaneous decomposition allows for a multidimensional description of the artifact as both a collection of parts and as a purposeful entity.

The Design Browser's conceptual framework has three parts: *building blocks* (i.e., a hierarchical decomposition of objects that comprise the system), *functional components* (i.e. sets of system objects grouped together with a recognizable purpose), and *system issues* (i.e., system information about important features such as application language or lines of source code). Building blocks are components of the software system which are manipulated by the functions. Issues add a high-level system view.

2.2 Case-Base

The Design Browser uses the conceptual framework as a method to organize and display information about several existing designs. This information is organized as cases, or more specifically, *stories* (Kolodner, 1993). Stories are interesting examples of previous designs that illustrate how existing systems implement the building blocks or functions, or how design issues are resolved.

The Design Browser has three complementary sets of features: building blocks, functions, and issues. Stories about design features are hierarchical as well as graphical explanations of the features. The hierarchical explanation makes the large amounts of data about previous designs more manageable and easily understood. The graphical view increases understanding of the component or function by providing an alternative to text-based description.

The stories about design features have three parts: generic description, examples, and illustrations. Figure 1 shows these parts in graphical form. The generic description is hierarchical and consists of a topic and associated details. The topic is a high-level description of the building block, function, or issue which has been generalized from several application-specific designs. The details provide lower-level, application-independent information about interesting aspects of the features (i.e., where a particular design of a feature differs from other designs), and is a direct link to the second part of the story -- examples. The examples consist of application-specific descriptions of the system features. They also contain 'outcome' information. Outcome information consists of opinions about the effectiveness of a particular instantiation of a feature.

The last part, an illustration, is a graphical representation that complements the text of a topic, detail, or example.

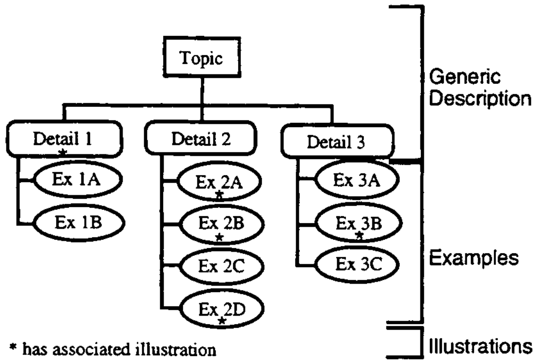


Figure 1. Structure of Design Browser's Stories

2.3 *Role of the Design Browser*

Software development often occurs in an environment with collaborative teams of people representing different organizational entities (e.g., each group may work for a different company or government organization). The Design Browser attempts to support three sets of people: the customer/user who specify the functionality of the system, the designer/developer who turn the user specifications into software, and management who is responsible for overall system effectiveness and functionality. The Design Browser allows the customer/user, designer/developer, and management to explore, compare, and contrast true differences necessitated by application requirements versus differences that occur serendipitously -- differences that potentially degrade reuse and increase development costs. The Design Browser's framework provides all participants with three tools: (1) a common frame of reference, i.e., a high level conceptual framework within which to consider the current design with respect to previous designs at a detailed, technical level; (2) a common vocabulary with which to articulate the current design to facilitate rapid comparison to previous designs; and (3) a more general context in which to explore both higher-level and management issues like reuse or choice of application software.

3. A DESIGN BROWSER FOR COMMAND MANAGEMENT SOFTWARE DESIGN

The Design Browser concept was implemented for NASA Goddard Space Flight Center's command management system (CMS) applications. For almost all CMS applications, significant reuse of previous command management system design concepts or software fails to occur at any level of abstraction.

The Design Browser CMSs proposes a unifying conceptual framework for the CMS software. The framework was developed by studying the documentation describing seven existing or 'in-process' CMSs. Its building blocks consist of CMS components common to all CMSs: data inputs, commands, activities, etc. Likewise its functions define the major functional requirements of command management: scheduling, command preprocessing, etc. Salient issues involved in CMS design and reuse were identified through conversations with NASA command management personnel (i.e., customer/user, designer/developer, NASA management).

4. EMPIRICAL EVALUATION OF THE CMS DESIGN BROWSER

A study was conducted to assess the effectiveness and perceived usefulness of the CMS Design Browser in answering the questions about reuse which arise before and at Critical Design Reviews. An important feature of the evaluation of the CMS Design Browser was that the study participants were actual personnel involved in the specification, development, and management of creating new command management systems. Since the CMS Design Browser was intended to support experienced CMS personnel, the strategies, performance, and opinions of experienced personnel was of specific interest.

The empirical evaluation of the proposed Design Browser for CMSs consists of both formal process and product evaluation and participant questionnaires. The formal process and product evaluation strives to validate the usefulness of the CMS Design Browser to experienced personnel. The process measures examine how the participants used the CMS Design Browser to arrive at a useful conclusion. The product measures assess the extent to which useful information could be extracted from the CMS Design Browser by the participants. The questionnaires sought the participants' perceived usefulness of the CMS Design Browser. The next sections describe the evaluation and questionnaire results.

4.1 *Evaluation*

Over two weeks, twelve participants used the CMS Design Browser to complete ten tasks. The first five tasks were conceptual tasks which demonstrated the ability of the CMS Design Browser to support the user in navigating to and viewing all features of the CMS Design Browser. The last five tasks were task-based activities which demonstrated the ability of the CMS Design Browser to support the participants in representative design tasks. The observations were recorded manually, captured by the CMS Design Browser in a computer file, and recorded on video tape. The computer file and video tape were used to augment any information that was missing or vague in the evaluator's notes. A total of 120 tasks were observed. The process of each task was compared to the expected process.

Process Measures. Each task was analyzed with respect to four process criteria: expected process, short process, long process, and different process. These classifications are explained below:

Expected process. The procedure that the participant used is in the "family" of processes that is expected by the designer of the CMS Design Browser. If the participant goes "browsing" in the middle or at the end of a task, that is not included when comparing the participant's process and the expected process. However, this is an interesting phenomenon and shows that the CMS Design Browser can support browsing and encourages users to browse.

Short process. The procedure is shorter or less involved than those in the "family" of processes that is expected by the designer of the CMS Design Browser. For example: For one of the tasks it is expected that the participant will go to a particular issue topic and look in detail at two of the associated examples. If a participant looks in detail at only one of the examples, then the participant's process is classified as a "short process."

Long process. The procedure is longer or more involved than those in the "family" of processes that is expected by the designer of the CMS Design Browser. A process is only "long" if the additional actions or steps can not be attributed to browsing. For example: For one of the tasks it is expected that the participant will look at one function topic and its first detail and its associated examples. If a participant reviews a different topic and then looks at the expected function topic and its associated detail and examples, then the participant's process is classified as a "long process." A process is a long process if it is longer than the expected process at any stage of the process (i.e., the beginning, the middle, or the end).

Different process. The participant's procedure is different than the process that was expected by the designer of the CMS Design Browser. That is the entire expected process can not be found in the participants' process. For example: For one of the tasks it is expected that the participant will look at two topics with their associated details and examples. If a participant looks at the same first topic but then goes to a different second topic and never looks at the expected second topic, then the participant's process is classified as a "different process."

Table 1 indicates the number of tasks of each type. One hundred and two of the tasks were the "expected process" and only eighteen tasks were mismatches (i.e., short process, long process, and different process).

Table 2 provides information to compare the user groups to each other. Each user group had similar results to the participants as a whole with the designer/developer group having the highest percentage of expected processes.

Product Measures. Data on each of the 120 tasks were analyzed with respect to four product measure criteria: match, partial match, no match, and no result. These classifications are explained below:

Match. The participant's result matches the expected result or is more involved than was expected. A participant may give a more involved or complete answer than the one expected when the participant is experienced and can draw more detailed conclusions from the information than is expected. For example: When the participants were asked to find a spacecraft which uses immediate commanding, some of the participants equated this with the similar ability to send real-time commands from locations other than the Payload Operations Control Center and listed two spacecraft instead of the one that was expected.

Table 1 Process Evaluation Results

Classification	Total (120)	Break down	Percent
Expected Process	102		85.00
Short	10		8.33
Skipped expected example		3	
Skipped outcome information		1	
Skipped expected details and examples		4	
Skipped expected topic		2	
Long	4		3.33
Looked at additional topic first		3	
Looked at additional topic in middle		1	
Different	4		3.33
Used different topic		3	
Used different second topic		1	

Table 2 Process Measures for Participation Groups

Classification	Managers		Customer		Developer	
	Number	Percent	Number	Percent	Number	Percent
Expected Process	33	82.5	32	80	37	92.5
Short Process	3	7.5	4	10	3	7.5
Long Process	2	5	2	5	0	0
Different Process	2	5	2	5	0	0

Partial match. The result partially matches the expected result. There are two types of partial match: the participant may leave one part of the result uncompleted or may have obtained a different result for one part.
No match. The result does not match the expected result.
No result. The participant does not believe an appropriate result can be found with the CMS Design Browser. The participant may or may not be looking at the correct information (i.e., they have used the same process), but for some reason the participant does not feel he/she can get an appropriate result from the CMS Design Browser.

Table 3 indicates the number of tasks of each type. One hundred and one tasks were "matches," only twenty-one tasks were mismatches (i.e., partial match, no match, and no result).

Table 4 shows the breakdown of the tasks by user group. Each user group had similar results to the participants as a whole with the manager group having the highest percentage of product matches.

Table 3 Product Evaluation Results

Classification	Total (120)	Break down	Percent
Match	101		84.17
Partial Match	16		13.33
Portion of result is missing		11	
Portion of result is different		5	
No Match	1		0.83
No Result	2		1.67

Table 4 Product Measures for Participation Groups

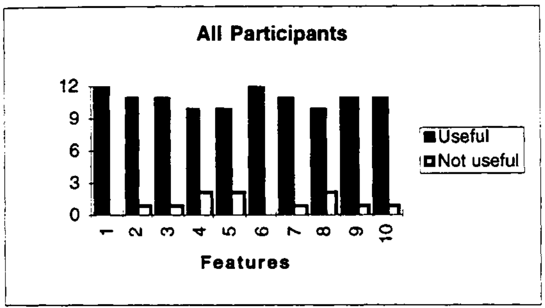
Classification	Managers		Customer		Developer	
	Number	Percent	Number	Percent	Number	Percent
Match	37	92.5	31	77.5	33	82.5
Partial Match	3	7.5	6	15	7	17.5
No Match	0	0	1	2.5	0	0
No Result	0	0	2	5	0	0

4.2 Questionnaires

Along with the objective data collected and analyzed above, there was also subjective data collected to determine the opinions of the participants as to the perceived usefulness of the CMS Design Browser. Even if the Design Browser can be used by the intended users, if the intended users do not perceive it as useful, it will not be used. This section describes the initial and final questionnaires and discusses the answers that were given by the participants.

Initial Questionnaire. The initial questionnaire asked the participants to give the number of years they had been involved in the design of command management systems and to rate their familiarity with the process. The mean number of years of experience for each group were very similar with the overall mean number of years which was 4.03 (minimum one year six months, maximum nine years and two months). All but three of the participants rated themselves as "Very familiar" or "Familiar" with the process of CMS design. Of the three who rated themselves lower, two said they were "Somewhat familiar" and one said "Not very familiar". However, the one participant who considered himself "Not very familiar" had three years of experience on two different missions and appeared knowledgeable while using the CMS Design Browser to complete the representative tasks.

Final Questionnaire. The final questionnaire asked about the usefulness of the CMS Design Browser and its various features. Figure 2 shows the responses. The CMS Design Browser was thought to be useful by all the participants and all of the features were considered useful by a majority of the participants.



1. CMS Design Browser overall
2. Generic descriptions
3. Individual examples
4. Illustrations
5. Outcome information
6. Menu bar and button navigation
7. Navigation Outline navigation
8. Simultaneous examples
9. Save conceptual design
10. Review saved conceptual design

Figure 2. Questionnaire Results

A second part of this questionnaire asked the participants to rank the features in terms of usefulness (1 = most useful, 9 = least useful). The results of this ranking are shown in Table 5.

Table 5: CMS Design Browser Feature Ranks

1	Individual examples
2	Simultaneous examples
3	Generic descriptions
4	Outcome information
5	Illustrations
6	Navigation Outline navigation
7	Review saved conceptual design
8	Save conceptual design
9	Menu bar and button navigation

4.3 Summary of Evaluation

Overall, the evaluation of the CMS Design Browser was quite positive. Data supports the conclusion that the system allows users to retrieve useful information in a predictable way. The CMS Design Browser was also perceived as a useful tool, particularly with respect to examples of previous designs. It appears that the Design Browser would make a good institutional memory for the design of command management systems.

ACKNOWLEDGEMENTS

This work is sponsored by NASA GSFC grant NAG 5-2226, William J. Potter, Technical Monitor.

REFERENCES

Chen, D. and P.J. Lee (1993). On the study of software reuse using reusable C++ components. In: *Journal of Systems and Software*. **20**, 19-36.

Curtis, B. (1989). Cognitive issues in reusing software artifacts. In: *Software Reusability Vol. 2*, (T.J. Biggerstaff and A.J. Perlis (Eds.)), 269-287. ACM Press, New York, NY.

Cusumano, M.A. (1991). *Japan's Software Factories*. Oxford University Press, New York, NY.

Domeshek, E.A. and J.L. Kolodner (1992). Toward a case-based aid for conceptual design. In: *International Journal of Expert Systems*. **4**(2), 201-220.

Domeshek, E.A., M.F. Herndon, A.W. Bennett, and J.L. Kolodner (1993). A case-based design aid for conceptual design of aircraft subsystems. In: *IEEE Conference on AI Applications '94*. in review.

Kolodner, J.L. (1993). *Case-Based Reasoning*. Morgan Kaufmann, San Mateo, CA.

Lanergan, R.G. and C.A. Grasso (1984). Software engineering with reusable designs and code. In: *IEEE Transactions on Software Engineering*. **10**(5), 498-501.

Mili, H., J. Witt, R. Radai, W. Wang, K. Strickland, C. Boldyreff, J. Heger, W. Scherr, L. Olsen, and P. Elzer (1994). Practitioner and softClass: a comparative study of two software reuse research projects. In: *Journal of Systems and Software*. **25**,147-170.

Neighbors, J.M. (1984). The Draco approach to constructing software from reusable components. In: *IEEE Transactions on Software Engineering*. **10**(5), 564-574.

Ockerman, J.J., C.M. Mitchell, and W.J. Potter (1993). A case-based design browser to aid human developers in reuse of previous design concepts. In: *Proceedings of 1994 IEEE International Conference on Systems, Man, and Cybernetics*, San Antonio, TX, 1757-1762.

Ockerman, J.J. (1994). *A Case-Based Design Browser to Aid Human Developers in Reuse of Previous Design Concepts*. Unpublished Master's Thesis, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA.

Parnas, D.L., P.C. Clements, and D.M. Weiss (1989). Enhancing reusability with information hiding. In: *Software Reusability Vol. 1*, (T.J. Biggerstaff and A.J. Perlis (Eds.)), 141-157. ACM Press, New York, NY.

Prieto-Diaz, R. (1993). Status report: software reusability. In: *IEEE Software*. **10**(3), 61-66.

Rasmussen, J. (1986) *Information Processing and Human-Machine Interaction*. North-Holland New York, NY.

A DESIGN METHODOLOGY FOR OPERATOR DISPLAYS OF HIGHLY AUTOMATED SUPERVISORY CONTROL SYSTEMS

David A. Thurman and Christine M. Mitchell

Center for Human-Machine Systems Research
School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0205
[dave, cm]@chmsr.isye.gatech.edu

Abstract: This research proposes an extension to the Operator Function Model (OFM) as the basis for a methodology to guide the design of interactive monitoring and control interfaces (i.e., operator interfaces), specifically designed to support the monitoring tasks of supervisory system control. The goal of such an interface is to engage the operator in the monitoring activity to ensure that fault detection occurs in an efficient and reliable way.

Keywords: human-machine interface, human-centered design, human supervisory control, computer interfaces, interaction mechanisms

1. WHY IS MONITORING IMPORTANT? WHY IS IT DIFFICULT?

With the increasing prevalence of computer automation in complex systems control, the role of the human operator has changed: humans predominantly monitor rather than control (Rasmussen & Rouse, 1981; Rouse, 1986; Sheridan, 1992). Unfortunately, humans are poor at monitoring (Wickens & Kessel, 1981). One proposed solution to this problem has been to automate the monitoring function. However, humans cannot be completely replaced as monitors of complex systems. Humans operators provide the ultimate safety system - they are flexible and opportunistic, able to respond to unforeseen events. The question is not how to replace the human operator as the monitor of the system, but rather how to make the operator more effective.

Humans are ineffective monitors of complex engineering systems for two interrelated reasons. The first concerns the *nature* of the monitoring task and the second relates to the *interfaces* provided for monitoring tasks. The monitoring task has two basic problems. First, monitoring is a boring, *passive* task that consists mainly of watching computer displays for changes in system state. Prior to the advent of supervisory control, the operator was *actively engaged* in monitoring and controlling the system because very few of those tasks were automated (Zuboff, 1991). With the introduction of automation, however, the operator is no longer actively involved. Instead, s/he is a passive observer, prone to distraction, inattentiveness, and complacency (Sheridan, 1992).

Second, monitoring is often an *unstructured*, ill-defined task. The monitoring requirements of a complex system containing thousands of sensors are overwhelming; however, operators are typically given little direction concerning the types of events for which they should monitor. Instead they are told 'watch everything'. As a result, the operator has no plan for surveying the system and could miss critical information about one component of the system simply because s/he forgot to check it.

Furthermore, the interfaces operators use to monitor do not support the task. Typically, operators monitor system status information. In many cases, rather than being designed, displays are simply developed by taking all available system status information and building a set of displays structured along a physical decomposition of system components. This hardware-centered approach (Billings, 1991) to interface development leads to an interface which does not support the tasks the operator needs to accomplish. The result is that the monitoring interface, and as a result, the monitoring task, is boring and passive. The interface does not engage the operator's attention. The monitoring task consists largely of watching a multitude of computer displays for changes in displayed information indicating the occurrence of a critical event. With such interfaces, monitoring is like watching a forest of tress in early autumn, waiting to see which tree will be the first to drop a leaf. It is no wonder that operators become inattentive when faced with such a boring task.

2. DESIGNING DISPLAYS AND INTERACTION

Research in the design of computer interfaces for complex systems has particular relevance to supporting effective monitoring, for the computer interface is the medium through which the operator monitors the system. Thus, an effective interface design can lead to effective monitoring, so long as that design is accomplished with the monitoring activity in mind. In particular, this research addresses the design of information displays to support the operator's monitoring activity and the design of human-computer interaction in the interface to actively engage the operator in the tasks which support the monitoring function.

2.1 Designing Information Content

There are several bodies of research that provide background and insights about the problems and challenges of designing interfaces to support monitoring and control functions for supervisory control systems. Hutchins and Norman's gulfs of evaluation and execution (Hutchins, Hollan, & Norman, 1986; Norman, 1986) delineate cognitive issues which display design for operators of complex systems must address. Woods' concept and ten year retrospective on representation aiding (Woods, 1991a; Woods, 1991b) provides an important backdrop for a discussion of interface design, particularly with respect to the semantics of the interface. Kirlik's ecological task analysis (Kirlik, Kossack, & Shively, 1994) offers a mechanism with which to analyze an operator's activities and examines the extent to which the perceptual aspects of the display support, or could support, fluent behavior. Finally, Mitchell's use of the operator function model (Mitchell, 1987; Mitchell & Saisi, 1987) demonstrates another strategy for designing displays in which the information content and level of detail are tailored to the user's needs via the specifications in the model of operator activity. See Thurman (1995) for a discussion of the impact of this research on the design of the information content of displays for monitoring.

2.2 Designing Interaction

Conventionally, monitoring is passive -- the operator just watches a screen. In reality, the operator is supposed to be checking and verifying the states of system components. To ensure that operators do indeed monitor, artificial task aids (e.g., checklists and sensor logs) are often introduced. Checklists require the operator to follow a set procedure for monitoring the system, executing a series of steps designed to ensure that the operator checks all critical components and subsystems. Sensor logs are another method, although slightly less restrictive, of ensuring that the operator monitors the system. The use of logs generally requires the operator to write down the readings of several critical state variables from each subsystem. Sensor logs may be used to provide long-range trending data, but their primary purpose is to ensure that the operator monitors the system (Bainbridge, 1987).

Unfortunately, while artificial tasks may provide some much needed structure to the operator's monitoring task, it is not clear that they actually improve monitoring effectiveness. Checklists shift the operator's attention away from the computer interface and may ironically result in the operator becoming more focused on completing the checklist than on carrying out the activities the checklist is designed to ensure are completed. While sensor logs may provide a gentler, less intrusive, way of introducing structure into the monitoring task, operators become quite proficient at writing down numbers without noticing what the numbers indicate (Bainbridge, 1987).

The strongest benefit of artificial tasks is that they introduce interaction between the operator and his/her environment. They provide action that helps keep the operator awake and engaged with the system. Without these artificial tasks, it is feared that operators will quickly become bored and complacent, further reducing their effectiveness.

Interaction between the operator and his/her environment, then, is a key ingredient in effective monitoring. This research explores the use of human-computer interaction to replace artificial tasks with an interactive monitoring and control interface which has the benefits (interaction, involvement, structure) of these tasks but not the drawbacks (distraction from the monitoring task). The goal is to transform monitoring interfaces from passive visual displays into interactive interfaces that use human-computer interaction to engage the operator in the monitoring task.

3. A DESIGN METHODOLOGY FOR INTERACTIVE MONITORING AND CONTROL INTERFACES

Figure 1 shows a schematic representation of the design methodology. The first step is to construct a model of the operator's monitoring and control function with a particular emphasis on the operator's monitoring activities. This model is further augmented through information flow analysis to represent the information requirements of each monitoring and control task and the flow of information between activities. The next step is to review the integrated operator function/information flow model and identify the interface components (i.e., information displays and interaction) to be used to supply the information to the operator to support each task. The last step is to design the information and interaction components of the interface using a set of design guidelines for interaction monitoring and control interfaces.

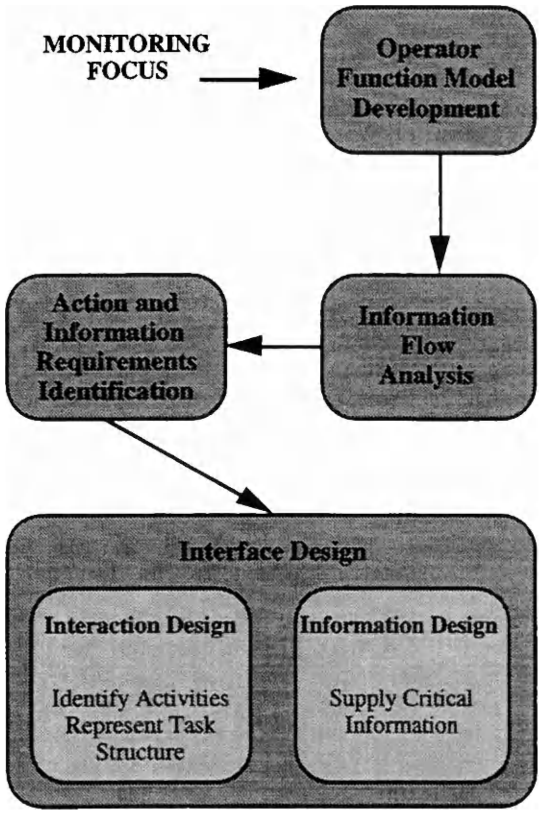


Fig. 1. Methodology for the design of interactive monitoring and control interfaces.

3.1 Modeling the Operator's Function

An OFM-based approach to display design focuses on normal operator decision making in operator functions such as monitoring and predictable fault detection and compensation (Mitchell & Saisi, 1987). The OFM is a hierarchic/heterarchic network in which major operator control functions constitute the highest, or heterarchical, level. The hierarchy is defined by operator subfunctions, control actions, control commands, and information requirements that the operator undertakes to carry out individual control functions (Mitchell, 1987).

Although the OFM has been shown to be very effective as a tool for the design of control interfaces (Mitchell & Saisi, 1987) it does not explicitly model monitoring activities. This research extends the OFM to the operator's monitoring function and demonstrates its use as a tool to support effective interface design for both monitoring and control. This extension is accomplished by explicitly representing the flow of information between operator activities and by analyzing monitoring not just as a singular activity but as an activity that is integrated into other activities.

3.2 Information Flow Analysis

One problem with monitoring is that the operator is often required to remember information obtained in one activity for use in another activity, the latter often occurring at a later point in time. External memory aids can help in this task, and operators often write down information or print out copies of displayed information that they know they will need at a later point in time. In other cases, operators access log files or other information stores from which they can retrieve the information when it is needed. These sorts of activities are identified and documented in the OFM by representing the flow of information between operator activities and the flow of information in and out of data stores.

Information flow analysis augments the operator function model with representations of the information required to support each monitoring and control activity. Traditional applications of the operator function model have indicated the system state information required to make control decisions. In this step of the proposed design methodology, particular attention is paid to the information requirements of the operator's monitoring activities and the flow of information among those activities. To accomplish this, it is useful to understand the types of monitoring activities that are likely to be modeled and the types of information flows and requirements that need to be represented. Thurman and Mitchell (1994) identify several different types of monitoring activities.

One role of the operator of a supervisory control system is to periodically check the state of the system, subsystems, components, etc. This responsibility includes checking the configuration of systems components (e.g., discrete states such as a valve position) as well as checking continuous output parameters (e.g., temperature). In this case, monitoring is effectively a singular activity independent of any control activities. Effectively supporting this type of monitoring requires identifying the information needs of the activity and presenting that information in the appropriate form and at the appropriate time.

A different type of monitoring takes place during control activities. Many systems require the operator to execute control procedures which consist of a number of steps. Each step must be individually monitored to ensure that it had the proper effect on the system and to determine the appropriate next step. Effectively supporting this activity requires identifying and representing the steps of the procedure, the opportunities for monitoring and control, and the relationships (order, branching, etc.) between the steps.

Monitoring for discrete events places an additional burden on the operator because it often requires the operator to store data gathered over a period of time and evaluate them at some later point. This process is usually aided by the use of external memory aids (checklists, sensor logs, etc.), but it still requires the operator to recall (from the memory aid) the information to perform an evaluation. Effectively supporting this activity requires identifying the information that needs to be collected over time, storing that information in the interface (possibly invisible to the operator), and then presenting the aggregated information (or performing the evaluation and presenting the result) to the operator in a timely manner.

3.3 Identify Information and Action Requirements

Support for the operator's monitoring and control function can be broken down into two sets of requirements: information requirements and action requirements. Information requirements consist largely of the system information required to complete an activity. Action requirements consist largely of the operator actions that need to be supported (i.e., afforded by the interface) to allow the operator to complete an activity. These action requirements specify how the operator gathers data, integrates raw data into abstract indications of system state, evaluates that system state against a target state, and determines the appropriate action to take based on that evaluation. Included in the action requirements is the operator's knowledge of what control actions exist within the system and how they are executed. The information necessary to support these requirements is found by analyzing the model of the operator's monitoring and control functions and identifying task structure (i.e., the action requirements) and system information requirements of each activity.

Each operator activity in the upper levels of the operator function model has a task structure. The task structure is drawn from the model's decomposition of that activity and specifies the steps of the activity and the operational information associated with that activity (i.e., how the task is performed). Control activities have a task structure that specifies a sequence of control actions and the methods of determining the next step after the conclusion of any step in the sequence. High-level monitoring activities have a task structure that specifies what should be monitored (i.e., which parameters, components, configuration, etc.) and how they should be monitored (i.e., what are the criteria, how are the parameters evaluated, etc.)

3.4 Interface Design

This research proposes that a computer interface, at its simplest decomposition, consists of two components: an information component and an interaction component. The prevalent approaches to interface design for complex systems concentrate exclusively on the information component and ignore the design of the interaction component. The use of human-computer interaction as a representation medium has been completely ignored.

This research proposes to use the interaction component to represent the action requirements of monitoring and control activities and to use the information component to supply the information required by those activities. That is, the interface is designed so that the interaction component affords completion of the required monitoring and control activities by providing the actions necessary to complete those activities, while the information component presents the system information necessary to complete those activities. Given this proposed mapping and the analysis of the operator's monitoring and control activities, the final step in the methodology is to design the interface to support those activities.

Interaction Design - Representing Tasks and Activities. The purpose of interaction design is to represent the tasks and activities the operator needs to complete in a way that makes it obvious to the operator what needs to be done and provides the tools necessary to complete those activities. The goal is to develop a more interactive interface which engages the operator's attention and makes the operator a more effective controller and monitor of the system. Avoiding the pitfall of adding to the operator's information processing burden requires a balance of interaction and information that is tailored to support the identified monitoring and control activities. Appropriate interactive interface technologies (e.g., command buttons, menus, scrollbars, drag-and-drop, etc.) can be used to identify critical activities to the operator, represent procedures, and support monitoring.

An interactive monitoring interface identifies critical activities, provides the means for completing those activities, and provides an indication of the current status of those activities. Identifying the activities and providing the means for completing them are part of the role of the interaction component. Integrating the identification of the critical activities together with the means for completing the activities reduces the complexity of the operator's interface manipulation task. A set of guidelines for incorporating interaction into the interface is discussed in (Thurman & Mitchell, 1994).

Control procedures involve some (often substantial) degree of monitoring and a second role of the interaction component is to explicate the monitoring and control actions required to complete a procedure. This research proposes the design of *integrated monitoring and control displays* to support the operator's execution of these types of procedures. An integrated monitoring and control display uses (1) interface interaction techniques to identify the sequence of steps (both control and monitoring) involved in a procedure and (2) visual information tokens to support the monitoring activities within the procedure.

The operator's responsibilities often include periodic scans of the system components looking for anything out of the ordinary. These monitoring activities have a repeatable pattern to them: read sensor, evaluate against expected value (either predetermined or based on current control procedures), analyze result. The interaction component can be designed to support this activity by allowing the operator to select the parameter of interest (through interaction, e.g., drag-and-drop) and automatically update a 'monitoring display' tailored to present the current value, limit conditions, past values, or any other information required to make the assessment.

Information Design - Supplying Critical Information. The purpose of the information component is to supply the information the operator needs to complete the tasks and activities represented by the interaction component. The goal is to supply the exact information the operator needs, without requiring the operator to cognitively manipulate the information to evaluate it.

Identifying critical activities and providing the means for completing them was identified as part of the role of the interaction component. Providing an indication of the current status of those activities, a high-level activity overview, is one role of the information component. This representation should indicate the current status of all critical activities and be visible during the execution of any of the critical activities.

In addition, although system monitoring procedures may specify how to monitor low-level parameters, operators are also interested in the overall state of the system as defined by a set of critical parameters or the state of each subsystem. The information component can be used to provide an indication of these parameters and/or subsystems.

Finally, each system monitoring activity requires certain information in order to be completed. This information may be directly available from system state data or it may require manipulation into a more abstract form to support the monitoring activity. Effectively supporting the operator's monitoring activities means tailoring the system data into the form of information required by the monitoring activities and providing it at the appropriate time in an easily understood form.

4. GT-TIMI: THE GEORGIA TECH TASK-INTERACTIVE MONITORING INTERFACE

Using the proposed methodology, a proof-of-concept monitoring and control interface was designed for the domain of NASA satellite ground control. The Georgia Tech Task-Interactive Monitoring Interface (GT-TIMI) supports more effective monitoring by (1) elucidating the monitoring requirements of the system, (2) providing a high-level overview of subsystem status, (3) providing displays tailored to the operator's monitoring and control activities, (4) communicating task structure (e.g., steps in a procedure) of the operator's activities using human-computer interaction, and (5) providing a high-level overview of the status of the activities required of the operator.

Figure 2 shows one of the displays developed for GT-TIMI. This display integrates the control actions required to retrieve scientific and engineering data from the spacecraft together with the information (and actions) necessary to monitor its progress. Other elements of the GT-TIMI interface are discussed in (Thurman & Mitchell, 1994) and (Thurman, 1995).

An empirical evaluation was conducted to compare the performance of operators using a conventional operator interface and the GT-TIMI interface. Experimental results indicate that use of the GT-TIMI interactive monitoring and control interface resulted in higher fault detection rates, lower fault detection times, fewer procedural errors, and a higher rate of completion of monitoring and control activities. This can be partially attributed to GT-TIMI allowing operators to complete control procecdure more rapidly and thus have more time to perform routine monitoring tasks. The experimental results are discussed in detail in (Thurman, 1995).

ACKNOWLEDGMENTS

This research was supported in part by NASA Goddard Space Flight Center grants NAG 5-2227 (A. William Stoffel, Technical Monitor) and NAG 5-1044 (Walt Truszkowski, Technical Monitor).

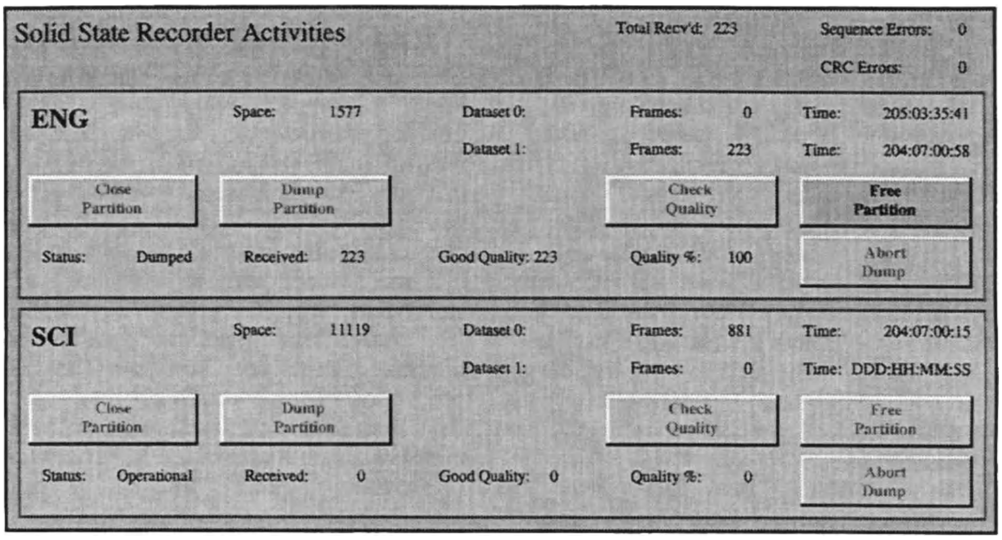


Fig. 2. Integrated Monitoring and Control Display from GT-TIMI

REFERENCES

Bainbridge, L. (1987). Ironies of automation. In: *New Technology and Human Error* (J. Rasmussen, K. Duncan, & J. Leplat (Ed.)), 271-283. John Wiley & Sons Ltd., New York.

Billings, C. E. (1991). *Human-Centered Aircraft Automation: A Concept and Guidelines* . 103885. NASA.

Hutchins, E. L., Hollan, J. D., & Norman, D. A. (1986). Direct manipulation interfaces. In: *User Centered System Design* (D. A. Norman & S. W. Draper (Ed.)), 87-124. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Kirlik, A., Kossack, M. F., & Shively, R. J. (1994). *Ecological task analysis: Supporting skill acquisition in dynamic interaction* . Unpublished manuscript. Center for Human-Machine Systems Research, School of Industrial and Systems Engineering, Georgia Institute of Technology.

Mitchell, C. M. (1987). GT-MSOCC: A domain for modeling human-computer interaction and aiding decision making in supervisory control systems. *IEEE Transactions on Systems, Man, and Cybernetics*, **SMC-17(4)**, 553-572.

Mitchell, C. M., & Saisi, D. L. (1987). Use of model-based qualitative icons and adaptive windows in workstations for supervisory control systems. *IEEE Transactions on Systems, Man, and Cybernetics*, **SMC-17(4)**, 573-593.

Norman, D. A. (1986). Cognitive engineering. In: *User Centered System Design* (D. A. Norman & S. W. Draper (Ed.)), 31-61. Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Rasmussen, J., & Rouse, W. B. (Ed.). (1981). *Human Detection and Diagnosis of System Failures.* Plenum Press, New York.

Rouse, W. B. (Ed.). (1986). *Advances in Human-Machine Systems Research.* JAI Press, Inc., Greenwich, CN.

Sheridan, T. B. (1992). *Telerobotics, Automation, and Human Supervisory Control.* The MIT Press, Cambridge, MA.

Thurman, D. A. (1995) *Improving Operator Effectiveness in Monitoring Complex Systems: A Methodology for the Design of Interactive Monitoring and Control Interfaces.* Unpublished Master's Thesis, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia.

Thurman, D. A., & Mitchell, C. M. (1994). A methodology for the design of interactive monitoring interfaces. In: *Proceedings of the1994 IEEE International Conference on Systems, Man, and Cybernetics*, 1738-1744. San Antonio, Texas.

Wickens, C. D., & Kessel, C. (1981). Failure detection in dynamic systems. In: *Human Detection and Diagnosis of System Failures* (J. Rasmussen & W. B. Rouse (Ed.)), 155-169. Plenum Press, New York.

Woods, D. D. (1991a). Computer based information displays: The theory and practice of representation design in the computer medium - An overview of representation aiding. In: *The Theory and Practice of Representation Design in the Computer Medium - Collected Working Papers as of September 1, 1991* (D. D. Woods (Ed.)), Cognitive Systems Engineering Laboratory, Department of Industrial and Systems Engineering, The Ohio State University, Columbus, OH.

Woods, D. D. (1991b). Representation aiding: A ten year retrospective. In: *Proceedings of the 1991 IEEE International Conference on Systems, Man, and Cybernetics*, 1173-1176. Charlottesville, VA.

Zuboff, S. (1991). *In the Age of the Smart Machine - The Future of Work and Power.* Basic Books, New York.

A DESIGNER'S ASSOCIATE: SOFTWARE DESIGN SUPPORT FOR COMMAND AND CONTROL SYSTEMS

John G. Morris and Christine M. Mitchell

Center for Human-Machine Systems Research
Georgia Institute of Technology
Atlanta, GA 30332-0205

Abstract: Design support for command and control software development represents a new, but interesting, aspect of the human-machine systems engineering problem. The design support approach described in this paper is rooted in a computer-based Designer's Associate which facilitates learning at the organizational level via the collection and diffusion of domain-specific experience.

Keywords: Computer aided design, command and control, systems design, system analysis

1. INTRODUCTION

As automated control systems proliferate in such applications as space, aviation, and manufacturing, the human role in such endeavors shifts from that of manual to supervisory controller (Sheridan, 1976). In a supervisory role, the operator is more concerned with cognitive rather than physical tasks. Such tasks emphasize monitoring, situation assessment, and failure management (Woods and Roth, 1988) rather than physical, manual, and/or clerical duties. With this change in emphasis has come a corresponding change in the role of the control system designer. The designer's, or more likely, the design team's, performance, or lack thereof, determines to a large degree the effectiveness of the real-time supervisory controller and hence the ultimate safety and effectiveness of the overall system. Rasmussen and Goodstein (1987) argue that, in contemporary, highly automated systems, supervisory control is not simply the task of the operator, but is, in actuality,

> ...a cooperative effort within a group consisting of designers, the automatic (computer-based) control system, and the operating staff. This three-way arrangement arises from the fact that the decisions of designers are embedded in the automatic system as well as the training of the operators (Rasmussen and Goodstein, 1987, p. 663).

With his or her ability to partially automate the control of a process, the designer affects a division of labor between human and machine and establishes their means of interaction (e.g., Rasmussen, 1986). In so doing, the designer becomes inexorably intertwined within the supervisory control of a process.

Given this state of affairs, the research described here embraces an expanded vision of supervisory control which encompasses both the human operator and the control system designer(s). It takes the view that *computer-based design support* represents a new, but interesting, aspect of human-machine systems engineering research. This view arises from three assumptions: 1) operator effectiveness hinges upon highly uncertain designer performance, 2) design support strategies based on the simple provision of salient human-factors guidelines are inadequate, and 3) given the rapidly developing information needs stemming from changes in personnel, design problems, and technology, a computer-based approach is essential.

A designer's performance, and thus the operator's effectiveness, is an uncertain thing, for while the designer's capability to field ever more powerful systems has increased rapidly, knowledge of how "best" to design and deploy such systems has not (Woods & Roth, 1988). Clearly, the designer has need of some form of *design support.* Traditional approaches to design support, however, hinge upon the provision of prescriptive recommendations for identifying and incorporating human factors into a product. According to Rasmussen, Pejtersen, and Goodstein (1994) design support of this sort is often inadequate, because it is domain independent in character and therefore lacks the familiar cues which would otherwise inform a highly context-dependent design process. In fact, this type of guidance tends to be ignored by experienced designers who lack human factors expertise. It is often only after problems arise that human factors guidelines are considered (Rasmussen *et al.*, 1994).

Even so, simply furnishing application-specific guidance (e.g., Mitchell, 1995) which contains the relevant cues may, nonetheless, fail to produce better designs (i.e., improved operator performance). In many cases, the expertise required to make use of such guidance is absent (Walz, Elam, and Curtis, 1993). That is, either the salience of prescriptive recommendations goes unnoticed or the design team is so struggling to produce *anything* (Brooks, 1975; Gibbs, 1994) that such recommendations appear irrelevant to the immediate problem.

Therefore, achieving *consistently effective design support* requires that one address an absence of domain-specific expertise as well as a dearth of human-factors expertise. This paper proposes a *computer-based Designer's Associate* which addresses these issues by focusing on the collection and diffusion of domain-specific experience. Such experience includes, among other things, examples of the successful application of human factors guidelines as well as illustrations of system performance inadequacies that may result from a failure to apply them or apply them appropriately. The overarching goal of the Designer's Associate is to facilitate learning at an organizational level with respect to the particulars of a given domain and to the utilization of human factors knowledge within that domain.

Although the Designer's Associate concept is thought to have broad applicability, it is embedded in the details of *satellite command and control systems*–principly in software design. Hence, this paper continues with a brief discussion of the problematic nature of software development. Software development is then contrasted with science–a more successful enterprise. The insight thus obtained strongly suggests a particular approach to design support. The description of the Designer's Associate that follows illustrates how this approach might be transformed into a computational architecture.

2. SOFTWARE DEVELOPMENT AND SCIENCE PRACTICE

After 50 years of continual refinement, software development continues to be problematic. Virtually every software system is crafted from scratch. There is essentially no codification of domain-specific experience. As a result, while successful solutions are often overlooked, mistakes continue to be repeated project after project (Gibbs, 1994).

It is posited that this state of affairs results from an absence of learning at an organizational level. In other words, what an individual learns over time seldom benefits more than himself or herself and his or her immediate team members. Diffusion of experience occurs only indirectly as a result of a designer's movement from project to project and organization to organization. In the absence of organizational learning, only limited, personal experience is left to both formulate plausible solutions and judge their adequacy. It is inevitable that progress under such circumstances will be slow, painful, and expensive.

With regard to the absence of learning, satellite command and control software development is no exception. Despite the development of many systems and a rather stable set of requirements, most problems are solved as if they were novel. Improvement typically results at an organizational level not from learning but from underlying technological change. The absence of organizational learning has a number of undesirable effects. Three are of particular interest in the domain of command and control software development. First, systems exhibit substantial variation both internally and externally (i.e., from a user point of view) in spite of similar fundamental requirements. Second, because recurring problems are rarely solved once and for all, new problems introduce an inordinate amount of technological risk. Finally and of much concern to NASA, productivity fails to increase.

Science, on the other hand, provides not only an example of organizational learning, but of its underpinnings as well. Scientific communities achieve a level of objectivity and a rate of knowledge growth which stands in stark contrast to the subjective nature of most software development and the near absence of knowledge growth that such development efforts exhibit. To be sure each science discipline has the luxury of defining its own problems and working on them till the solution is deemed satisfactory. However, though these factors may contribute to the success of scientific endeavors, they are by no means its principle cause. The objectivity and rapid progress of science stems from several, much more important, factors: 1) the public character of its methods, 2) an aversion to misunderstandings, and 3) a research tradition that is firmly based on one or more past scientific achievements.

The public character of science has two principle impacts. First, it provides a foundation for scientific objectivity. In a public forum of journals and conferences, the soundness of one's ideas are not determined by the individual's subjective judgment, but by the community in which he or she works. Furthermore, the final arbiter of controversy is experience. Not *private* experience, but *public* experience in the form of observations or experiments which, if one takes the trouble, can be repeated (Popper, 1970).

Second, the public character of science has a significant impact on the rate at which innovations are adopted. Because the members of a particular scientific community are in communication with one another, the probability that a given member of this community will adopt an innovation increases over time. Thus, for a time, knowledge growth increases at an exponential rate. In the absence of communication, the probability that a particular innovation will be adopted remains constant; hence, knowledge grows in a linear fashion (Crane, 1972).

The impact of science's public character is enhanced by the fact that each scientific community strives to avoid misunderstanding. Ideas are expressed in one and the same language. Moreover, the members of a scientific community express their theories in a form that can be tested (Popper, 1970). Thus, communication is made as efficient and effective as possible.

Finally, the men and women who work in a particular field prepare themselves by studying their field's past achievements. They learn the relevant concepts, laws, and theories within the context of their applications rather than in the abstract. Members of a scientific community so trained are committed to the same rules and standards and rarely disagree over fundamentals. The individual scientist can take the achievements of the field for granted and begin work where they leave off rather than build the field anew (Kuhn, 1970).

In comparison, software development, to some, might appear the antithesis of a science. First, the design team, in contrast to the individual scientist, works in virtual isolation. Hence, judgment is often based on untested assumptions; knowledge growth is slow and halting; and innovation when it occurs most often results not from the diffusion of domain-specific knowledge, but from exogenous influences (e.g., technological change).

Second, because the design team works in isolation, the language each team uses to describe its problem and proposed solutions remains fragmented in local forms (Ockerman & Mitchell, 1994). Moreover, the language adopted by a given team is directed not at the larger community, but at meeting the needs of the design team alone. Descriptions thus contrived often obscure similarities and exaggerate differences between problems that have much in common.

Finally, there are typically no past domain-specific achievements which are accepted by most as the foundation for further practice. In the absence of public discourse it is virtually certain that no such achievement will be forthcoming. If by chance this happens not to be the case, it is unlikely that an achievement of this magnitude will be recognized as such.

In short, an increase in individual learning, should it occur, does not automatically result in an increase in organizational learning. Knowledge fuels learning only if communication is continuous (Rycroft & Kash, 1994). Thus, it is the public character of science and the persistence of its acquired knowledge that provide insight into the attributes of consistently effective design support. Such design support must not only facilitate the transfer of knowledge across *domain* boundaries (e.g., between human factors research and satellite command and control practice), it must facilitate the spanning of *spatial* boundaries (i.e., design teams) and *temporal* boundaries (i.e., project start and end dates) as well.

3. THE DESIGNER'S ASSOCIATE

The Designer's Associate serves as a focal point for communication across domain, spatial, and temporal boundaries. The concept is grounded in the assumption that the most fundamental prerequisite for organizational learning is the long-term persistence and accessibility of domain-specific knowledge. The Designer's Associate represents a proposal to ensure persistence by embedding domain-specific knowledge in a computer-based medium (rather than rely solely upon the memory of human designers) in such a manner that typical barriers to communication (i.e., domain, spatial, and temporal boundaries) largely disappear.

Therefore, the foundation of organizational learning and the communication upon which it relies is a persistent base of empirical knowledge in which relevant domain independent and application-specific guidance is embedded. This base of knowledge persists and grows despite the ceaseless formation and dissolution of design teams and the perpetual starting and ending of projects. Thus, the Designer's Associate is directed primarily at enhancing the collection and diffusion of design relevant information. The designer uses the Designer's Associate to both gather information relevant to the creation of a new system and all of its attendant design artifacts as well as to construct the design artifacts themselves.

Within the Designer's Associate, empirical knowledge is taken to be: 1) *design artifacts* such as requirements specifications, detailed design descriptions, operations concepts, etc., 2) any associated commentary on these artifacts, and 3) domain-specific standards, rules, generalizations, theories, etc. Empirical knowledge as well as guidance originating from outside the domain is organized along several dimensions: 1) part-whole decomposition, 2) means-end relations, and 3) knowledge category.

In addition, the constituent parts of this base of knowledge may be linked through a network of references residing in the descriptions of various entities (e.g., a statement of requirements may reference another entity containing a set of human factors guidelines). This organizational structure and network of references together constitute a *map of the design territory* for a given domain (e.g., Rasmussen *et al.*, 1994). Therefore, the Designer's Associate raises the structural characteristics of design knowledge to a level that is explicit and subject to manipulation and exploration.

Part-whole decompositions and means-end relations are used to organize information about a particular design effort. Because they are hierarchical in nature, these organizational strategies allow a designer to explore the design territory at a level of detail appropriate to his or her needs. Means-end relations provide a mechanism for linking goals to implementation details. Part-whole decompositions, on the other hand, provide a mechanism for understanding a particular entity as either 1) a monolithic whole without consideration of its constituent parts or 2) in terms of its parts and their relationships. Thus, the description of a given entity may consist of a number of elements. One describes the whole and makes reference to the constituent parts. Other elements describe each of the parts and make reference to any subparts. And so on.

Each level of a means-end hierarchy (e.g., Rasmussen, 1986) describes the same system using a different set of principles. Descriptions at one level of the hierarchy appear as constraints at lower levels. As one moves up through a means-end hierarchy, one obtains a deeper understanding of the goals a system is intended to achieve. Whereas, by moving down through the hierarchy, one acquires a more detailed

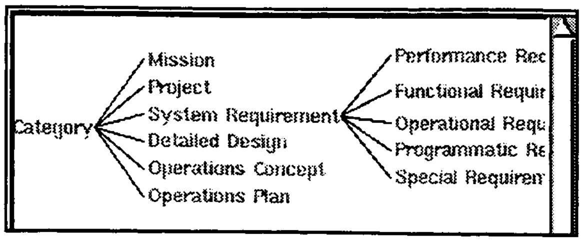


Fig. 1. Partial Category Structure for Satellite Command and Control Domain. Each category in this display is a hypertext link.

explanation of how these goals are accomplished. Means-end relations link the constituent parts of one design artifact description to that of another (e.g., individual requirements are linked to specific detailed design elements via means-end relations). In the course of design, one transforms information at one level to information at another.

Whereas part-whole decompositions and means-end relations serve to organize information about a particular design effort, knowledge categories provide a mechanism for organizing information about a design domain. Categories play a particularly important role, for they serve to organize large amounts of information in ways that require little cognitive effort (Rosch, 1988). For example, knowledge categories of particular interest in the realm of satellite command and control system design might include: system requirement, detail design specification, operations plan, etc. In addition, the system requirement category may have several subcategories that represent more specialized classes of requirement. Figure 1 illustrates a partial category hierarchy for satellite command and control.

The members of a category can be whole design artifacts or their constituent parts. General knowledge is also stored in the category dimension. Such knowledge might refer to: 1) subcategory structure, 2) category members, 3) attributes of category members, 4) exemplars, 5) generalizations about the part-whole decompositions of category members, 6) generalizations about the means-end relations of category members, and 7) human factors or human-computer interaction guidelines that are particularly salient with respect to category members.

The web of relationships (some of which fall outside of the three organizational dimensions) that exists between various entities are referenced within the descriptions of the these entities. Each of these references is a link within the map of the design territory. The links provide the designer with a vehicle for moving from one entity to another. Each description has a textual component in which references and graphics may be embedded. Each category member also has a number of derived references that pertain to

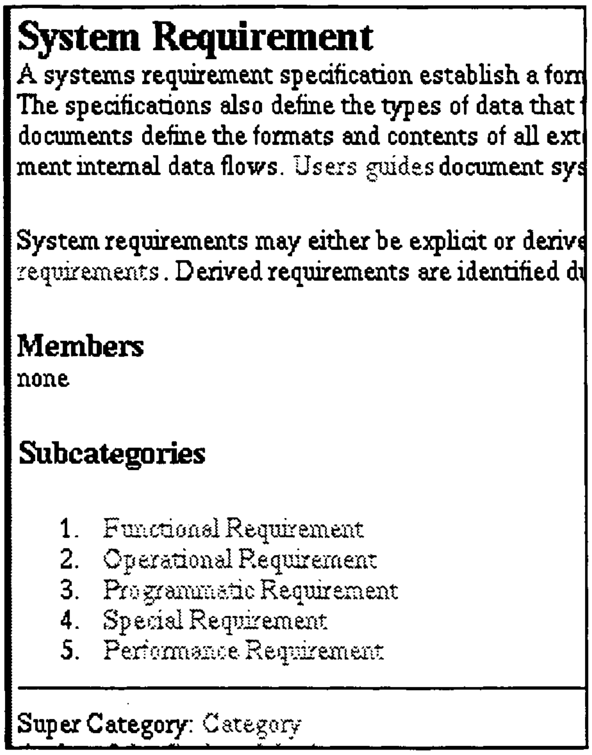


Fig. 2. Fragment of a Category Definition Page. This category has no members, but does have several subcategories. It is a subcategory of the highest level category–*Category*. Words shown in gray are hypertext links.

its category membership, its part-whole composition, and its means-ends relations. Category definitions, however, contain a list of references to category members, a list of references to subcategories, and a reference to the category in which the given category is a subcategory.

The Designer's Associate maintains its base of knowledge in an Oracle® database. It has the appearance of a World Wide Web-like browser (e.g., Mosaic, Netscape®). Highlighted hypertext links allow one to navigate within the design territory. Figure 2 illustrates a category definition page. This page describes the *system requirement* category which is also shown in Figure 1 in the context of a category hierarchy. Figures 1 and 2 also illustrate the two primary methods used by the Designer's Associate to display information. Trees are used to show structure (e.g, category structure, part-whole composition, and means-end relationship); whereas, principally textual descriptions emphasize non-structural information. Nevertheless, predominately textual descriptions contain implicit references to the structural characteristics of the local context.

The descriptions of categories and their members are derived in part based on the contents of the database. For example, the subcategories of *system requirement* are determined via a database query. The basic definition of a category such as *system requirement* may not change over time, but its members and subcategories, most likely, will. In terms of interaction, this means that the act of creating a category, for example, alters the displayed representation of its supercategory without resort to an editing operation.

The Designer's Associate serves both as a browsing tool and as an authoring tool. Descriptions may be edited; category membership can be altered; new categories can be defined, etc. More importantly, however, the components of existing design artifacts can be incorporated into new designs. By combining both a browser and an authoring tool, the Designer's Associate provides mechanisms for both the collection and diffusion of relevant information.

4. ORGANIZATIONAL STRUCTURE RATIONALE

The structure of the Designer's Associate knowledge base is grounded in several assumptions. First, experienced designers often reuse or adapt past solutions when faced with problems similar to those encountered in the past (e.g., Klein, 1993; Klein, 1989; Adelson and Soloway, 1988, Adelson and Soloway, 1985). Second, Klein's assertion that decision making in real-world settings is the result of situation assessment and mental simulation (Klein, 1993) is valid within a design context. Third, one of the most effective ways to assist designers is to "support in some way their intuitive explorations of the space of solutions (Rasmussen *et al.*, 1994, p. 170)." Finally, designers in the course of their work perform four generalized tasks: learning (Walz *et al.*, 1993), problem decomposition (Goel & Pirolli, 1992), incremental elaboration (Goel & Pirolli, 1992), and transformation of end to means (Rasmussen *et al.*, 1994; Goel & Pirolli, 1992).

Klein's (1993) assertion that situation assessment is used to generate a plausible course of action entails, within the context of design, understanding how the current problem is similar to past problems and thus how the current problem is amenable to past solutions. Empirical knowledge organized as part-whole decompositions and means-end relations provides concrete examples of past decompositions, incremental elaboration, and information transformation. Hence, the organization of the knowledge base reflects the most general character of design tasks (i.e., decomposition, composition, transformation from end to means, incremental elaboration). By embedding such knowledge in a category structure recognition of similarity is significantly enhanced. In addition, by organizing knowledge in a map-like structure both learning and exploration are facilitated (Rasmussen *et al.*, 1994).

5. CONCLUSION

Computer-based design support represents a new, but interesting facet of human-machine system engineering research. More traditional approaches to design support grounded in prescriptive recommendations may fail to provide adequate guidance. It is not, however, that such recommendations lack salience. Rather, recommendations of this sort often lack the familiar cues which might otherwise signal their relevance to the designer inexperience in applying human factors considerations to his or her design. Moreover, a given design team may face so many problems that human factors guidelines appear irrelevant to them.

The Designer's Associate concept represents an attempt to achieve *consistently effective design support* by addressing both an absence of domain-specific expertise and a dearth of experience in applying human-factors guidelines. The Designer's Associate fulfills this goal by focusing on the task of facilitating learning at an organizational level with respect to the particulars of a given domain and the utilization of human factors knowledge within that domain. This task is accomplished by focusing on the collection and diffusion of domain-specific experience. The overarching goal of the research is to demonstrate the effectiveness of the chosen knowledge organizational structures and their method of presentation in facilitating organizational learning.

ACKNOWLEDGMENTS

This work was supported by NASA Goddard Space Flight Center grant NAG 5-2226, William J. Potter, Technical Monitor.

REFERENCES

Adelson, B. & Soloway, E. (1985). The role of domain knowledge in software design. *IEEE Transactions on Software Engineering*, **11**, 1351-1360.

Adelson, B. & Soloway, E. (1988). A model of software design. In: *The Nature of Expertise* (M. T. H. Chi, R. Glaser, & M. J. Farr (Ed.)), 185-208. Lawrence Erlbaum Associates, Publishers, Hillsdale, NJ.

Brooks, F. P. (1975). *The Mythical Man-Month: Essays on Software Engineering*. Addison-Wesley Publishing Company, Reading, MA.

Crane, D. (1972). *Invisible Colleges*. The University of Chicago Press, Chicago.

Goel, V. & Pirolli, P. (1992). The structure of design problem spaces. *Cognitive Science*, **16**, 395-429.

Klein, G. A. (1993). A recognition-primed decision (RPD) model of rapid decision making. In: *Decision Making in Action: Models and Methods* (G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Ed.)), 138-147. Ablex Publishing Corporation, Norwood, NJ.

Kuhn, T. S. (1970). *The Structure of Scientific Revolutions*. The University of Chicago Press, Chicago.

Mitchell, C. M. (1995). Cognitive engineering models: A prerequisite to the design of human-computer interaction in complex dynamic systems. To appear in: *Human-computer Interface Design: Success Cases, Emerging Methods, and Real-world Context* (P. Polson (Ed.)).

Ockerman, J. J., Mitchell, C. M., & Potter, W. J. (1994). A case-based design browser to aid human developers in reuse of previous design concepts. *Proceedings of 1994 IEEE International Conference on Systems, Man, and Cybernetics*, San Antonio, TX, 1757-1762.

Popper, K. (1970). The Sociology of knowledge. In: *The Sociology of Knowledge* (J. E. Curtis & J. W. Petras (Ed.)), 649-660. Praeger Publishers, New York.

Rasmussen, J. (1986). *Information Processing and Human-Machine Interaction: An Approach to Cognitive Engineering*. North-Holland, New York.

Rasmussen, J., Pejtersen, A. M., & Goodstein, L. P. (1994). *Cognitive Systems Engineering*. John Wiley & Sons, Inc., New York.

Rasmussen, J. & Goodstein, L. P. (1987). Decision support in supervisory control of high-risk industrial systems. *Automatica*, **23**, 663-671.

Rosch, E. (1988). Principles of categorization. In: *Readings in Cognitive Science* (A. Collins & E. E. Smith (Ed.)), 312-322. Morgan Kaufmann Publishers, Inc., San Mateo, CA

Rycroft, R. W. & Kash, D. E. (1994). Technology policy in a complex world. *Technology in Society*. **16**(3).

Sheridan, T. B. (1976). Preview of models of the human monitor/supervisor. In: *Monitoring Behavior and Supervisory Control* (T. B. Sheridan & G. Johannsen (Ed.)), 175-180. Plenum Press, New York.

Walz, D. B., Elam, J. J., & Curtis, B. (1993). Inside a software design team: Knowledge acquisition, sharing and integration. *Communications of the ACM*, **36**, 63-77.

Woods, D. D. & Roth, E. M. (1988). Cognitive systems engineering. In: *Handbook of Human-computer Interaction* (M. Helander (Ed.)), 3-44. Elsevier Science Publishing Co., Inc., New York.

AUTHOR INDEX