Handbook of Philosophical Logic 15

Dov M. Gabbay
Franz Guenthner *Editors*

Handbook of Philosophical Logic

Second Edition

Springer

HANDBOOK OF PHILOSOPHICAL LOGIC
2ND EDITION

VOLUME 15

HANDBOOK OF PHILOSOPHICAL LOGIC

2nd Edition

Volume 15

edited by D.M. Gabbay and F. Guenthner

Volume 1 – ISBN 0-7923-7018-X
Volume 2 – ISBN 0-7923-7126-7
Volume 3 – ISBN 0-7923-7160-7
Volume 4 – ISBN 1-4020-0139-8
Volume 5 – ISBN 1-4020-0235-1
Volume 6 – ISBN 1-4020-0583-0
Volume 7 – ISBN 1-4020-0599-7
Volume 8 – ISBN 1-4020-0665-9
Volume 9 – ISBN 1-4020-0699-3
Volume 10 – ISBN 1-4020-1644-1
Volume 11 – ISBN 1-4020-1966-1
Volume 12 – ISBN 1-4020-3091-6
Volume 13 – ISBN 978-1-4020-3520-3
Volume 14 – ISBN 978-1-4020-6323-7

HANDBOOK OF PHILOSOPHICAL LOGIC

2nd EDITION

VOLUME 15

Edited by

D.M. GABBAY
King's College, London, U.K.

and

F. GUENTHNER
Centrum für Informations- und Sprachverarbeitung,
Ludwig-Maximilians-Universität München, Germany

Springer

Editors
Prof. Dr. Dov M. Gabbay
King's College London
Dept. Computer Science
London
United Kingdom
dov.gabbay@kcl.ac.uk

Franz Guenthner
University of Munich
Centrüm für Informations- und Sprachverarbeitung
Oettingenstr. 67
80538 Munich
Germany
gue@cis.uni-muenchen.de

ISBN 978-94-007-0484-8 e-ISBN 978-94-007-0485-5
DOI 10.1007/978-94-007-0485-5
Springer Dordrecht Heidelberg London New York

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

CONTENTS

Editorial Preface — vii
Dov M. Gabbay

Lambda Calculi: A Guide — 1
Chris Hankin

Interpolation and Definability — 67
Dov M. Gabbay and Larisa L. Maksimova

Discourse Representation Theory — 125
Hans Kamp, Josef van Genabith and Uwe Reyle

Index — 395

PREFACE TO THE SECOND EDITION

It is with great pleasure that we are presenting to the community the second edition of this extraordinary handbook. It has been over 15 years since the publication of the first edition and there have been great changes in the landscape of philosophical logic since then.

The first edition has proved invaluable to generations of students and researchers in formal philosophy and language, as well as to consumers of logic in many applied areas. The main logic article in the Encyclopaedia Britannica 1999 has described the first edition as 'the best starting point for exploring any of the topics in logic'. We are confident that the second edition will prove to be just as good!

The first edition was the second handbook published for the logic community. It followed the North Holland one volume *Handbook of Mathematical Logic*, published in 1977, edited by the late Jon Barwise. The four volume *Handbook of Philosophical Logic*, published 1983–1989 came at a fortunate temporal junction at the evolution of logic. This was the time when logic was gaining ground in computer science and artificial intelligence circles.

These areas were under increasing commercial pressure to provide devices which help and/or replace the human in his daily activity. This pressure required the use of logic in the modelling of human activity and organisation on the one hand and to provide the theoretical basis for the computer program constructs on the other. The result was that the *Handbook of Philosophical Logic*, which covered most of the areas needed from logic for these active communities, became their bible.

The increased demand for philosophical logic from computer science and artificial intelligence and computational linguistics accelerated the development of the subject directly and indirectly. It directly pushed research forward, stimulated by the needs of applications. New logic areas became established and old areas were enriched and expanded. At the same time, it socially provided employment for generations of logicians residing in computer science, linguistics and electrical engineering departments which of course helped keep the logic community thriving. In addition to that, it so happens (perhaps not by accident) that many of the Handbook contributors became active in these application areas and took their place as time passed on, among the most famous leading figures of applied philosophical logic of our times. Today we have a handbook with a most extraordinary collection of famous people as authors!

The table below will give our readers an idea of the landscape of logic and its relation to computer science and formal language and artificial intelligence. It shows that the first edition is very close to the mark of what was needed. Two topics were not included in the first edition, even though

D. Gabbay and F. Guenthner (eds.),
Handbook of Philosophical Logic, Volume 15, vii–ix.
DOI 10.1007/978-94-007-0485-5, © Springer Science+Business Media B.V. 2011

they were extensively discussed by all authors in a 3-day Handbook meeting. These are:

- a chapter on non-monotonic logic
- a chapter on combinatory logic and λ-calculus

We felt at the time (1979) that non-monotonic logic was not ready for a chapter yet and that combinatory logic and λ-calculus was too far removed.[1] Non-monotonic logic is now a very major area of philosophical logic, alongside default logics, labelled deductive systems, fibring logics, multi-dimensional, multimodal and substructural logics. Intensive re-examinations of fragments of classical logic have produced fresh insights, including at time decision procedures and equivalence with non-classical systems.

Perhaps the most impressive achievement of philosophical logic as arising in the past decade has been the effective negotiation of research partnerships with fallacy theory, informal logic and argumentation theory, attested to by the Amsterdam Conference in Logic and Argumentation in 1995, and the two Bonn Conferences in Practical Reasoning in 1996 and 1997.

These subjects are becoming more and more useful in agent theory and intelligent and reactive databases.

Finally, fifteen years after the start of the Handbook project, I would like to take this opportunity to put forward my current views about logic in computer science, computational linguistics and artificial intelligence. In the early 1980s the perception of the role of logic in computer science was that of a specification and reasoning tool and that of a basis for possibly neat computer languages. The computer scientist was manipulating data structures and the use of logic was one of his options.

My own view at the time was that there was an opportunity for logic to play a key role in computer science and to exchange benefits with this rich and important application area and thus enhance its own evolution. The relationship between logic and computer science was perceived as very much like the relationship of applied mathematics to physics and engineering. Applied mathematics evolves through its use as an essential tool, and so we hoped for logic. Today my view has changed. As computer science and artificial intelligence deal more and more with distributed and interactive systems, processes, concurrency, agents, causes, transitions, communication and control (to name a few), the researcher in this area is having more and more in common with the traditional philosopher who has been analysing

[1] I am really sorry, in hindsight, about the omission of the non-monotonic logic chapter. I wonder how the subject would have developed, if the AI research community had had a theoretical model, in the form of a chapter, to look at. Perhaps the area would have developed in a more streamlined way!

such questions for centuries (unrestricted by the capabilities of any hardware).

The principles governing the interaction of several processes, for example, are abstract an similar to principles governing the cooperation of two large organisation. A detailed rule based effective but rigid bureaucracy is very much similar to a complex computer program handling and manipulating data. My guess is that the principles underlying one are very much the same as those underlying the other.

I believe the day is not far away in the future when the computer scientist will wake up one morning with the realisation that he is actually a kind of formal philosopher!

The projected number of volumes for this Handbook is about 18. The subject has evolved and its areas have become interrelated to such an extent that it no longer makes sense to dedicate volumes to topics. However, the volumes do follow some natural groupings of chapters.

I would like to thank our authors are readers for their contributions and their commitment in making this Handbook a success. Thanks also to our publication administrator Mrs J. Spurr for her usual dedication and excellence and to Kluwer Academic Publishers for their continuing support for the Handbook.

Dov M. Gabbay
King's College London

Logic	IT			
	Natural language processing	**Program control specification, verification, concurrency**	**Artificial intelligence**	**Logic programming**
Temporal logic	Expressive power of tense operators. Temporal indices. Separation of past from future	Expressive power for recurrent events. Specification of temporal control. Decision problems. Model checking.	Planning. Time dependent data. Event calculus. Persistence through time—the Frame Problem. Temporal query language. temporal transactions.	Extension of Horn clause with time capability. Event calculus. Temporal logic programming.
Modal logic. Multi-modal logics	generalised quantifiers	Action logic	Belief revision. Inferential databases	Negation by failure and modality
Algorithmic proof	Discourse representation. Direct computation on linguistic input	New logics. Generic theorem provers	General theory of reasoning. Non-monotonic systems	Procedural approach to logic
Non-monotonic reasoning	Resolving ambiguities. Machine translation. Document classification. Relevance theory	Loop checking. Non-monotonic decisions about loops. Faults in systems.	Intrinsic logical discipline for AI. Evolving and communicating databases	Negation by failure. Deductive databases
Probabilistic and fuzzy logic	logical analysis of language	Real time systems	Expert systems. Machine learning	Semantics for logic programs
Intuitionistic logic	Quantifiers in logic	Constructive reasoning and proof theory about specification design	Intuitionistic logic is a better logical basis than classical logic	Horn clause logic is really intuitionistic. Extension of logic programming languages
Set theory, higher-order logic, λ-calculus, types	Montague semantics. Situation semantics	Non-well-founded sets	Hereditary finite predicates	λ-calculus extension to logic programs

Imperative vs. declarative languages	Database theory	Complexity theory	Agent theory	Special comments: A look to the future
Temporal logic as a declarative programming language. The changing past in databases. The imperative future	Temporal databases and temporal transactions	Complexity questions of decision procedures of the logics involved	An essential component	Temporal systems are becoming more and more sophisticated and extensively applied
Dynamic logic	Database updates and action logic	Ditto	Possible actions	Multimodal logics are on the rise. Quantification and context becoming very active
Types. Term rewrite systems. Abstract interpretation	Abduction, relevance	Ditto	Agent's implementation rely on proof theory.	
	Inferential databases. Non-monotonic coding of databases	Ditto	Agent's reasoning is non-monotonic	A major area now. Important for formalising practical reasoning
	Fuzzy and probabilistic data	Ditto	Connection with decision theory	Major area now
Semantics for programming languages. Martin-Löf theories	Database transactions. Inductive learning	Ditto	Agents constructive reasoning	Still a major central alternative to classical logic
Semantics for programming languages. Abstract interpretation. Domain recursion theory.		Ditto		More central than ever!

Classical logic. Classical fragments	Basic background language	Program synthesis	A basic tool	
Labelled deductive systems	Extremely useful in modelling		A unifying framework. Context theory.	Annotated logic programs
Resource and substructural logics	Lambek calculus		Truth maintenance systems	
Fibring and combining logics	Dynamic syntax	Modules. Combining languages	Logics of space and time	Combining features
Fallacy theory				
Logical Dynamics	Widely applied here			
Argumentation theory games		Game semantics gaining ground		
Object level/ metalevel			Extensively used in AI	
Mechanisms: Abduction, default relevance			ditto	
Connection with neural nets				
Time-action-revision models			ditto	

	Relational databases	Logical complexity classes	The workhorse of logic	The study of fragments is very active and promising.
	Labelling allows for context and control.		Essential tool.	The new unifying framework for logics
Linear logic			Agents have limited resources	
	Linked databases. Reactive databases		Agents are built up of various fibred mechanisms	The notion of self-fibring allows for self-reference
				Fallacies are really valid modes of reasoning in the right context.
			Potentially applicable	A dynamic view of logic
				On the rise in all areas of applied logic. Promises a great future
			Important feature of agents	Always central in all areas
			Very important for agents	Becoming part of the notion of a logic
				Of great importance to the future. Just starting
			A new theory of logical agent	A new kind of model

CHRIS HANKIN

LAMBDA CALCULI: A GUIDE

1 INTRODUCTION

One of the universal notions of programming languages is functional abstraction. The methods of Java and the functions defined and used in functional programming languages, such as Haskell, are instances of this general notion. The inspiration for this form of abstraction mechanism comes from Mathematical Logic; notably Church's λ(lambda)-calculi and Schönfinkel's and Curry's Combinatory Logic. A proper study of these foundations leads to a better understanding of some of the fundamental issues in Computer Science. Areas in which they have had a major influence include:

Programming Language Design: We have already suggested the link with the notion of functional abstraction in programming languages. In addition, many of the typing notions found in modern programming languages have been inspired by the typing mechanisms found in these formal calculi. A notable example of this, to which we shall return, is the style of *polymorphism* which is employed in modern functional programming languages.

Semantics: One of the predominant schools of thought on this topic is *denotational semantics*. In this approach a typed λ-calculus is used as the meta-language; the meaning of a program is expressed by mapping it into a corresponding λ-calculus object. Understanding what such objects are requires that we should have a *model* of the calculus; the construction of such models has been the motivation for the subject of *domain theory*.

Computability: A classical use of the λ-calculus was in the study of computability; the study of the theoretical limitations of formal systems for describing computations. Indeed the first result in computability was a result concerning the relationship between the λ-calculus and Kleene's Recursive Functions. The (un-)decidability results familiar from automata theory have their analogues in the theory of the λ-calculus.

Natural Language Understanding: Since Richard Montague's pioneering use of λ-abstractions in natural language semantics in the 1970s there has been extensive use of extended calculi within the linguistics community.

I hope that this chapter will serve as an introduction to the λ-calculus for students of these areas.

D. Gabbay and F. Guenthner (eds.),
Handbook of Philosophical Logic, Volume 15, 1–66.
DOI 10.1007/978-94-007-0485-5_1,

2 SYNTAX

Classically, in set theory, a function is represented by its *graph*. The graph of a function defines a function by its input/output behaviour; for example, a unary function is represented by a set of pairs where the first component of each pair specifies the argument and the second component specifies the corresponding result. From this perspective, the function on pairs of natural numbers which adds its two arguments is represented as:

$$\{((0,0),0),((0,1),1),\ldots,((1,0),1),((1,1),2),\ldots\}$$

or:

$$\{((m,n),p) \mid m,n \in Num, p = n+m\}$$

Two functions are equal if they have the same graph; this notion of equality is referred to as *extensional* equality and we will return to it later.

From the point of view of Computer Science, this representation is not very useful. We are usually as interested in *how* a function computes its answer as in *what* it computes. For example, all sorting functions have the same graph and are thus (extensionally) equal but a significant part of the Computer Science literature has been devoted to the definition and analysis of different sorting algorithms, so we are clearly missing something. The casual use of the word "algorithm" in the last sentence is the key; we should represent a function by a rule, which describes how the result is calculated, rather than its graph. In this scheme, two functions are equal if they are both defined by the same (or equivalent) rules; this form of equality is called *intensional* equality. The λ-calculus[1] provides a formalism for expressing functions as *rules of correspondence* between arguments and results.

The λ-calculus consists of a notation for expressing rules, λ-notation, and a set of axioms and rules which tell us how to compute with terms expressed in the notation. A BNF specification of the λ-notation is:

$$\begin{array}{lcll}
<\lambda\text{-term}> & ::= & <\text{variable}> \mid & \\
 & & (\lambda <\text{variable}>< \lambda\text{-term}>) \mid & \textbf{(abs)} \\
 & & (< \lambda\text{-term}>< \lambda\text{-term}>) & \textbf{(app)} \\
<\text{variable}> & ::= & x \mid y \mid z \ldots &
\end{array}$$

It is more usual to present the syntax of a formal calculus using an inductive definition. The λ-calculus may defined in this style in the following

[1] There is a wide variety of different λ-calculi. The calculi differ along many axes: syntax, typing, rules of inference,.... When we talk of <u>the</u> λ-calculus we generally mean the pure, type-free $\lambda K\beta$-calculus which is the primary object of study in Barendregt's encyclopaedic book.

way. The class of λ-terms consists of words constructed from the following alphabet:

$x, y, z, \ldots$	variables
λ	
$(,)$	parentheses

We define terms formally as follows:

DEFINITION 1 (λ-terms).
The class Λ of λ-terms is the least class satisfying the following:

1. $x \in \Lambda$, x a variable
2. if $M \in \Lambda$ then $(\lambda x M) \in \Lambda$
3. if $M, N \in \Lambda$ then $(MN) \in \Lambda$

We specify the least class to avoid including "junk" terms; the clauses of the definition say what has to be in the class, not what shouldn't be. Some λ-terms are:

$$x \qquad (xz) \qquad ((xz)(yz)) \qquad (\lambda x(\lambda y(\lambda z((xz)(yz)))))$$

The intuition is that terms matching clause 2 correspond to function definitions, where the variable after the λ specifies the name of the formal parameter, and terms matching clause 3 correspond to function applications. Thus:

$$(\lambda xx)$$

should be compared to:

```
(\ x -> x)
```

in Haskell, or to:

```
public int id(int x) { return x;}
```

in Java[2]. To avoid the proliferation of parentheses, we will generally use an alternative notation for terms constructed according to clause 2 of the definition:

$$\lambda x.M$$

[2]The λ-term is type-free. This is in contrast to the Haskell function which is polymorphic and the Java method which is strongly (monomorphically) typed. Both the λ-term and the Haskell program are actually equivalent to a whole set of Java methods with an element for every type.

and moreover, we will elide internal λs and "."s and assume that abstraction associates to the right so that the following terms are equivalent:

$$\lambda x_1 \ldots x_n.M \;\equiv\; \lambda \vec{x}.M \;\equiv\; (\lambda x_1(\ldots(\lambda x_n M)\ldots))$$

where $\vec{x}$ is our notation for the sequence $x_1, \ldots, x_n$. We will generally use the symbol $\equiv$ to denote syntactic equality between terms.

The symbol λ acts as a variable binder in a similar fashion to $\int \ldots dx$ in integral calculus and the quantifiers $\exists$ and $\forall$ in predicate calculus. The set of bound variables is defined inductively by the following function, $BV : \Lambda \to \mathcal{P}(Var)$[3]:

$$\begin{array}{lcl} BV\,x & = & \emptyset \\ BV\,(\lambda x M) & = & (BV\,M) \cup \{x\} \\ BV\,(MN) & = & (BV\,M) \cup (BV\,N) \end{array}$$

We will also need the set of free variables in a term; these are defined inductively by the following function, $FV : \Lambda \to \mathcal{P}(Var)$:

$$\begin{array}{lcl} FV\,x & = & \{x\} \\ FV\,(\lambda x M) & = & (FV\,M) - \{x\} \\ FV\,(MN) & = & (FV\,M) \cup (FV\,N) \end{array}$$

When $(FV\,M)$ is the empty set, $\emptyset$, M is said to be *closed*; closed terms are sometimes called *combinators* and the class of all such terms is Λ^0. Notice that the sets of bound and free variables are not necessarily disjoint; x occurs both bound and free in:

$$x(\lambda xy.x)$$

Terms which are defined by clause 3 of the definition correspond to applications. We adopt the convention that application is left associative. Consequently:

$$MN_1 \ldots N_n \;\equiv\; M\vec{N} \;\equiv (\ldots(MN_1)\ldots N_n)$$

In the following, we also make use of the notion of subterm. A *subterm* of a λ-term is some part of the term which is itself a well-formed λ-term; we can generate the set of subterms using the function, $Sub : \Lambda \to \mathcal{P}(\Lambda)$, defined as follows:

$$\begin{array}{lcl} Sub\,x & = & \{x\} \\ Sub\,(\lambda x M) & = & (Sub\,M) \cup \{(\lambda x M)\} \\ Sub\,(MN) & = & (Sub\,M) \cup (Sub\,N) \cup \{(MN)\} \end{array}$$

[3] $f : A \to B$ means that f is a function which takes arguments from the "set" A to results in B. The notation $\mathcal{P}(A)$ constructs the powerset of A.

Although we have presented an recursive definition of subterms, we could think of "subterm" as being a reflexive, transitive, binary relation on terms. Our definition does not distinguish between different occurrences of the same subterm; to do so we would need to construct a multi-set of subterms but we will not require this refinement in the following.

Often, we will need the notion of a partially specified term, that is a term with "holes" in it. Such a term gives a context into which we can put other terms (to fill the holes!). The ability to construct contexts will clarify some definitions and generalise some results (for example see the generalisation of the Substitution Lemma later). We give an inductive definition of contexts for λ-terms:

DEFINITION 2 (Contexts).
The class $\mathcal{C}[]$ of λ-contexts is the least class satisfying:

1. $x \in \mathcal{C}[]$
2. $[] \in \mathcal{C}[]$
3. if $C_1[], C_2[] \in \mathcal{C}[]$ then $(C_1[]C_2[]), (\lambda x C_1[]) \in \mathcal{C}[]$

Notice that a hole is represented by $[]$. An example of a context is:

$$((\lambda x.[]x)M)$$

We will often give a name to a context, say $C[]$ for the one above, such names will always terminate with "[]". To represent the term generated by filling the holes in a context with some term, we write the name of the context with the term that is to fill the hole appearing between the square brackets. Thus:

$$C[\lambda y.y]$$

is the term:

$$(\lambda x.(\lambda y.y)x)M$$

Of course, a context may have many holes but they will all be filled with the same term; we could generalise this by labelling holes, in which case different holes could be filled by different terms, but we will not need such generality in here. Notice that variables in $FV(M)$ might become bound in $C[M]$ if the context has holes inside λ-abstractions.

A major limitation of the notation seems to be that we can only define unary functions; we introduce one formal argument at a time. The fact that this is not a real restriction was first observed by Schönfinkel. Given some binary function denoted by an expression in formal arguments x and y, say $f(x, y)$, then we define:

$$a \equiv (\lambda y(\lambda x(f(x, y))))$$

then a is equivalent to the original function but takes its arguments one at a time[4].

We next introduce the basic theory of equality between λ-terms. We will identify a set of "canonical" terms – the *normal* forms. We present material on reduction; this concerns how we compute with terms. Next we turn to semantics and abstractly characterise models of the λ-calculus. We then consider the relationship between the λ-calculus and other notions of computable functions. Finally we present a number of typed calculi.

3 THE BASIC THEORY

3.1 The theory λ

We can construct formulae from the terms; a theory then establishes certain formulae as axioms and provides rules of inference which enable us to derive new formulae. The true formulae (either axioms or formulae that can be derived from the rules) are called *theorems.*

We now present a theory of equality (or *convertibility*) between λ-terms. There are a number of reasonable requirements for such a theory:

1. An application term should be equal to the result obtained by applying the function part of the term to the argument. For example, suppose that Java methods can be higher-order (take methods as arguments and produce them as results) and that we have defined a higher-order variant of `id`. Then:

   ```
   id(fun)
   ```

   should surely be the same method as `fun` (for any appropriate method parameter `fun`).

2. Equality should be an equivalence relation.

3. Equal terms should be equal in any context.

These requirements go some way to motivating the theory λ which is shown in Figure 1.

The rule (ξ) is sometimes called the rule of weak extensionality. The rule (β) is the rule which corresponds to function application. The notation $M[x := N]$ should be read "replace free occurrences of x in M by N" (some care must be taken — we return to this later). The classical presentation of the theory also includes an α-rule which allows a change of bound variable names. Computer Science readers should compare the rule (β) to

[4]A function such as a, which takes its arguments one at a time, is often called a *curried* function (in honour of the logician Haskell B. Curry).

$$(\lambda x.M)N = M[x := N] \qquad (\beta)$$

$$M = M$$

$$\frac{M = N}{N = M}$$

$$\frac{M = N \quad N = L}{M = L}$$

$$\frac{M = N}{MZ = NZ}$$

$$\frac{M = N}{ZM = ZN}$$

$$\frac{M = N}{\lambda x.M = \lambda x.N} \qquad (\xi)$$

Figure 1. The theory λ

their intuitive understanding of the meaning of procedure calls in a familiar programming language.

We write:

$$\lambda \vdash M = N$$

to mean that $M = N$ is a theorem of λ and read the theorem as "M and N are convertible". The notation of λ-terms and this theory are variously called the λ-calculus (the name that we will use in the following), the $\lambda\beta$-calculus, the λK-calculus or the $\lambda K\beta$-calculus.

Note that:

$$M \equiv N \Rightarrow M = N$$

but:

$$\neg(M = N \Rightarrow M \equiv N)$$

For example:

$$(\lambda x.x)y = y$$

but the two terms are not identical.

Finally, we illustrate the use of the theory to prove a fundamental theorem, the Fixed Point Theorem. This will play an important role when we consider computability. Fixed points give meaning to self-referential constructs such as recursive method definitions or functions. The theorem fixes a term F and then states that there is another term which is a fixed point for F. The proof of the theorem is constructive in that it shows precisely how to construct the required term.

THEOREM 3 (The Fixed Point Theorem).

$$\forall F \in \Lambda, \exists X \in \Lambda . FX = X$$

Proof.
Let $W \equiv \lambda x.F(xx)$ and $X \equiv WW$. Then

$$X \equiv WW \equiv (\lambda x.F(xx))W = F(WW) \equiv FX$$

■

In a more familiar context, for example, 1 is a fixed point of the squaring function. The Fixed Point Theorem may seem quite surprising at first sight; it says that all terms have fixed points. For some terms, such as:

$$\lambda x.x$$

which is the identity function, this is obvious (all terms are fixed points of the identity!) but for others, such as:

$$\lambda xy.xy$$

it is not so clear. However, since the proof of the Fixed Point Theorem is constructive; it gives a recipe for constructing a fixed point of any term. In the second case above this leads to the following construction:

$$W \equiv \lambda x.(\lambda xy.xy)(xx) = \lambda x.\lambda y.(xx)y \equiv \lambda xy.(xx)y$$

The required fixed point is thus

$$(\lambda xy.(xx)y)(\lambda xy.(xx)y)$$

we can check that this is indeed a fixed point of the original term:

$$\begin{aligned}
&(\lambda xy.(xx)y)(\lambda xy.(xx)y) \\
&= \lambda y.((\lambda xy.(xx)y)(\lambda xy.(xx)y))y \\
&= (\lambda xy.xy)((\lambda xy.(xx)y)(\lambda xy.(xx)y))
\end{aligned}$$

The fixed point constructed for the identity function is:

$$(\lambda x.xx)(\lambda x.xx)$$

This term plays a special role in the theory which we shall return to later[5].

[5] For those readers familiar with domain theory, this term plays the same role as $\perp$. It is the least fixed point of the identity function (and many others!).

Fixed points are important in Computer Science. They play a fundamental role in the semantics of recursive definitions. For example, the factorial function:

$$\begin{array}{lcl} fac\ 0 & = & 1 \\ fac\ (succ\ n) & = & (succ\ n) \times (fac\ n) \end{array}$$

is a fixed point of the term:

$$\lambda fn.if(= n\ 0)\ 1\ (\times\ n\ (f(pred\ n)))$$

(of course we must be careful about reading too much into this term – $0, 1, \times$ are just formal symbols, variables, they have no deeper significance in the λ-calculus which we have defined so far). We shall return to this point later.

3.2 Substitution

We now return to the substitution operation used in the rule (β). A naive approach to defining this operation leads to the problem of "variable capture". This problem occurs when we naively substitute a term containing a free variable into a scope where the variable becomes bound. For example:

$$(\lambda xy.yx)y \neq \lambda y.yy$$

The free occurrence of y in the left hand term is analogous to a global variable in programming, in the right hand side the global variable has become confused with the bound variable (formal parameter). We will consider three different approaches to this problem before selecting one for use in the rest of this article.

Three Approaches

The Classical Approach: The first approach is based on Church's original treatment of substitution. We use the following definition:

1. $x[x := N] \equiv N$
2. $y[x := N] \equiv y$, if x is not the same as y
3. $(\lambda x.M)[x := N] \equiv \lambda x.M$
4. $(\lambda y.M)[x := N] \equiv \lambda y.M[x := N]$, if $x \notin FV\,M$ or $y \notin FV\,N$
5. $(\lambda y.M)[x := N] \equiv \lambda z.(M[y := z])[x := N]$, if $x \in FV\,M$ and $y \in FV\,N$, z a new variable
6. $(M_1 M_2)[x := N] \equiv (M_1[x := N])(M_2[x := N])$

We consider the three rules 3 to 5 in a bit more detail. Rule 3 applies when the variable being substituted for is bound at the outermost level; in this case there will be no free occurrences of x in the remainder of the expression and thus the substitution has no effect. Rule 4 is applicable when variable capture cannot occur, either x does not occur free in the body (in which case the substitution is a no-operation again) or the variable that is bound in the outermost level does not occur free in the term being substituted (no capture); in either case the substitution can be pushed through the λ to apply to the body. Rule 5 applies when variable capture could occur, that is when some substitution does take place and the variable bound at the outermost level does occur free in the term being substituted; in this case, we first rename the bound variable to a completely new variable.

Rule 5 is only valid under the assumption that terms which are similar, having the same free variables and differing only in their bound variables, are essentially the same. This is reasonable if we think about programming languages:

```
public int id(int y) { return y;}
```

the above method is clearly the same as the earlier one with the same name; we have only changed the formal parameters. In Church's original presentation of the λ-calculus there were two additional axioms; (α) formalises the above discussion and (η) introduces extensional equality (see below). The alpha rule is:

$$\lambda x.M = \lambda y.M[x := y], y \notin FV\,M \qquad (\alpha)$$

The Variable Convention: For our second definition of the substitution operation, which is introduced in Barendregt's book, we start with two definitions:

DEFINITION 4 (Change of Bound Variables).
M' is produced from M by a *change of bound variables* if $M \equiv C[\lambda x.N]$ and $M' \equiv C[\lambda y.(N[x := y])]$ where y does not occur at all in N and $C[]$ is a context with one hole.

DEFINITION 5 (α-congruence).
M is α-congruent to N, written $M \equiv_\alpha N$, if N results from M by a series of changes of bound variable.

According to the second definition, we have:

$$\lambda x.xy \equiv_\alpha \lambda z.zy$$

but not:

$$\lambda x.xy \equiv_\alpha \lambda y.yy$$

Notice that the first two terms are also equal by the rule (α) but the second two are not. Our strategy for defining substitution is as follows:

1. Identify α-congruent terms
2. Consider a λ-term as a representative of its equivalence class
3. Interpret $M[x := N]$ as an operation on the equivalence classes, using representatives according to the following *variable convention*:

DEFINITION 6 (Variable Convention).
If $M_1, \ldots, M_n$ occur in a certain context then in these terms all bound variables are chosen to be different from free variables[6].

With this strategy, we can define substitution as follows:

1. $x[x := N] \equiv N$
2. $y[x := N] \equiv y$, if $x \not\equiv y$
3. $(\lambda y.M)[x := N] \equiv \lambda y.(M[x := N])$
4. $(M_1M_2)[x := N] \equiv (M_1[x := N])(M_2[x := N])$

The variable capture problem has disappeared! — the reason for this is that for y to appear free in N in the context:

$$(\lambda y.M)[x := N]$$

would breach the variable convention so we would have to use a different representative of the α-equivalence class of $\lambda y.M$ (this is precisely what rule 5 in the classical approach makes explicit). In the following, we will adopt this convention and definition of substitution because it is easier to work with (there are less cases to consider in proofs). An example of its use is illustrated below:

$$\begin{array}{rcll}(\lambda xyz.xzy)(\lambda xz.x) & = & \lambda yz.(\lambda xw.x)zy & \text{by the variable convention} \\ & = & \lambda yz.(\lambda w.z)y & \\ & = & \lambda yz.z & \end{array}$$

However, before continuing we consider a third alternative.

[6]We have already implicitly employed this convention in the proof of the Fixed Point Theorem — consider what happens if x occurs free in the term F.

The de Bruijn Notation: The third approach to defining substitution avoids the problem of variable capture by banishing free variables. We revise the definition of λ-terms so that parameters occurring in the body of a term are referred to by natural numbers which uniquely identify the binding λ. For example:

$$\lambda.\lambda.2 \text{ is equivalent to } \lambda xy.x$$

This is the notation invented by de Bruijn and used in the Automath project, an automated theorem proving system. More formally the terms in de Bruijn's notation are defined inductively as the least set such that:

1. any natural number (greater than zero) is a term
2. If M and N are terms, then (MN) is a term
3. If M is a term, (λM) is a term

and (β) is replaced by:

$$(\lambda P)Q = P[1 := Q]$$

where:

$$\begin{array}{lcl}
n[m := N] & \equiv & \begin{cases} n & \text{if } n < m \\ n-1 & \text{if } n > m \\ rename_{n,1}(N) & \text{if } n = m \end{cases} \\
(M_1M_2)[m := N] & \equiv & (M_1[m := N])(M_2[m := N]) \\
(\lambda M)[m := N] & \equiv & \lambda(M[m+1 := N])
\end{array}$$

and

$$\begin{array}{lcl}
rename_{m,i}(j) & \equiv & \begin{cases} j & \text{if } j < i \\ j+m-1 & \text{if } j \geq i \end{cases} \\
rename_{m,i}(N_1N_2) & \equiv & rename_{m,i}(N_1)rename_{m,i}(N_2) \\
rename_{m,i}(\lambda N) & \equiv & \lambda(rename_{m,i+1}(N))
\end{array}$$

The following example illustrates the effect of this new β rule:

EXAMPLE 7.

$$\begin{array}{rcl}
\lambda.(\lambda.\lambda.2)1 & = & \lambda.(\lambda.2)[1 := 1] \\
 & \equiv & \lambda.\lambda.2[2 := 1] \\
 & \equiv & \lambda.\lambda.rename_{2,1}(1) \\
 & \equiv & \lambda.\lambda.2
\end{array}$$

(c.f. $(\lambda x.(\lambda yz.y)x)$).

Notice the rôle that *rename* takes in relabelling variable indices. There is a simple translation between standard λ-terms and de Bruijn terms (notice that α-congruent terms are equal in the de Bruijn notation):

$$\begin{array}{lcl}
DB\,x\ (x_1,\ldots,x_n) & = & i, \text{ if } i \text{ is the minimum such that } x \equiv x_i \\
DB\,(\lambda xM)\ (x_1,\ldots,x_n) & = & \lambda(DB\,M\ (x,x_1,\ldots,x_n)) \\
DB\,(MN)\ (x_1,\ldots,x_n) & = & (DB\,M\ (x_1,\ldots,x_n))(DB\,N\ (x_1,\ldots,x_n))
\end{array}$$

The de Bruijn notation is not very readable but the beta rule is easy to implement; indeed it inspired the Categorical Abstract Machine — an efficient mechanism for the implementation of functional languages.

The Substitution Lemma

From now on, we will assume the variable convention unless otherwise stated.
We now present a result which allows us to reorder substitutions, The Substitution Lemma.

LEMMA 8 (The Substitution Lemma).
If x and y are distinct variables and $x \notin FV\,L$ then

$$M[x := N][y := L] \equiv M[y := L][x := N[y := L]]$$

The proof is a straightforward induction on the structure of M.

Substitution has a number of other useful properties with respect to convertibility:

1. $M = M' \Rightarrow M[x := N] = M'[x := N]$
2. $N = N' \Rightarrow M[x := N] = M[x := N']$
3. $M = M', N = N' \Rightarrow M[x := N] = M'[x := N']$

These properties are useful but care should be taken when applying them. A major property of functional languages is *referential transparency*; the property which allows equals to be substituted by equals. The properties of substitution appear to be related to this concept but referential transparency is more. For example, the following inference does not follow from (1) to (3):

$$N = N' \Rightarrow \lambda x.x(\lambda y.N) = \lambda x.x(\lambda y.N')$$

This is because we cannot express the two sides of the second equality in the correct form:

$$\lambda x.x(\lambda y.N) \text{ is not the same as } (\lambda x.x(\lambda y.z))[z := N]$$

since N may contain free occurrences of y. The correct formulation of the property of referential transparency, also referred to as Leibniz Law, is:

LEMMA 9 (Referential Transparency).
Let $C[]$ be a context, then

$$N = N' \Rightarrow C[N] = C[N']$$

The proof is by induction on the structure of contexts. The variable convention takes care that we do not inadvertently capture any variables in this substitution.

3.3 Extensionality

The convertibility relationship, $=$, is intensional equality; two terms are equal if they encode the same algorithm in some sense. This does not equate some terms which we might naturally think of as equal. For example, consider a term which has one bound variable and applies some constant term (i.e. a term that does not contain free occurrences of the bound variable) to any term bound to the variable:

$$\lambda x.Mx$$

this term should surely be equal to M since if we apply either $\lambda x.Mx$ or M to some term N, we end up with MN. The formula:

$$\lambda x.Mx = M$$

is not a theorem of λ; there are two ways we can extend λ to make the above formula a theorem. Firstly, we could add a new rule to the theory, giving the new theory $\lambda + ext$:

$$\frac{Mx = Nx}{M = N} \qquad x \notin (FV\,MN) \qquad\qquad (\mathbf{ext})$$

Alternatively, we can add a new axiom, giving the new theory $\lambda\eta$ (as proposed by Church):

$$\lambda x.Mx = M, x \notin FV\,M \qquad (\eta)$$

In fact, the following result can be shown:

LEMMA 10. $\lambda + \mathbf{ext}$ *and* $\lambda\eta$ *are equivalent*

The calculus based on $\lambda\eta$ or $\lambda + \mathbf{ext}$ is alternatively called the $\lambda\eta$-calculus, the $\lambda\beta\eta$-calculus, the $\lambda K\eta$-calculus or the $\lambda K\beta\eta$-calculus. Practically, from the point of view of functional programming, the $\lambda\eta$-calculus is not as important as the $\lambda\beta$-calculus since the rule (η) is not normally implemented. The term $\lambda x.Mx$ is a weak head normal form and is thus distinguishable from M; the former is a "value" whilst the latter may lead to a non-terminating computation. Even in an eager language, such as Standard ML, the two terms are distinguished. However, the $\lambda\eta$-calculus does have some theoretical significance which we shall return to later.

3.4 Consistency and Completeness

For a theory to be useful, there must be some theorems and not all closed formulae should be theorems. The former is satisfied provided that the theory has at least one axiom. The latter is slightly trickier and is quite a fragile property; a theory which satisfies this constraint is called *consistent*. Both of the theories presented here are consistent but it is very easy to lose consistency as we shall see.

We start by formalising the concept. First, some definitions:

DEFINITION 11. An *equation* is a formula of the form:

$$M = N$$

where $M, N \in \Lambda$.

DEFINITION 12. An equation is *closed* if $M, N \in \Lambda^0$.

DEFINITION 13 (Consistency).
If $\mathcal{T}$ is a theory with equations as formulae then $\mathcal{T}$ is *consistent*, written $\mathrm{Con}(\mathcal{T})$, if it does not prove every closed equation.
If $\mathcal{T}$ is a set of equations then $\lambda + \mathcal{T}$ is formed by adding the equations of $\mathcal{T}$ as axioms to λ. $\mathcal{T}$ is consistent, also written $\mathrm{Con}(\mathcal{T})$, if $\mathrm{Con}(\lambda + \mathcal{T})$.

Both of the theories that we have dealt with in this Section, λ and $\lambda\eta$ are consistent (see Barendregt's book). The property of consistency can be lost by adding a single equation. We define the following three terms:

$$\begin{array}{lcl} \mathbf{S} & \equiv & \lambda xyz.xz(yz) \\ \mathbf{K} & \equiv & \lambda xy.x \\ \mathbf{I} & \equiv & \lambda x.x \end{array}$$

Notice that:

$$\begin{array}{lcl} \mathbf{S}MNO & = & MO(NO) \text{ by three applications of } (\beta) \\ \mathbf{K}MN & = & M \text{ by two applications of } (\beta) \\ \mathbf{I}M & = & M \end{array}$$

Now, if we add the equation:

$$\mathbf{S} = \mathbf{K}$$

to either λ or $\lambda\eta$ we get an inconsistent theory. This can be proved as follows (we elide some of the steps):

EXAMPLE 14.
$$\begin{array}{rcl} \mathbf{S} = \mathbf{K} & \Rightarrow & \mathbf{S}ABC = \mathbf{K}ABC \text{ for all } A, B, C \\ & \Rightarrow & AC(BC) = AC \end{array}$$
Now consider the case when $A = C = \mathbf{I}$, then since $\mathbf{I}A = A$ for all A:
$AC(BC) = AC \Rightarrow B(\mathbf{I}) = \mathbf{I}$
Now consider the case when $B = \mathbf{K}D$ for some arbitrary D, then:
$B(\mathbf{I}) = \mathbf{I} \Rightarrow D = \mathbf{I}$
and thus, since D was arbitrary, all terms are equal to the constant term $\mathbf{I}$.

Consideration of the foregoing motivates the following definition:

DEFINITION 15 (Incompatibility).
Let $M, N \in \Lambda$, then M and N are *incompatible*, written $M \# N$, if $\neg Con(M = N)$.

We now turn to the notion of completeness. Yet again, we start by making some definitions:

DEFINITION 16 (Normal Forms).
If $M \in \Lambda$, then M is a *β-normal form*, written β-nf or nf, if M has no subterms of the form $(\lambda x.R)S$
If $M \in \Lambda$, then M *has a* β-nf if there exists an N such that $N = M$ and N is a β-nf.

Some (non-)examples of normal forms:

$\lambda x.x$ is a nf
$(\lambda xy.x)(\lambda x.x)$ has $\lambda yx.x$ as a nf
$(\lambda x.xx)(\lambda x.xx)$ does not have a nf

By analogy, a $\beta\eta$-nf is a β-nf which also does not contain any subterms of the form:

$$(\lambda x.Rx) \text{ with } x \notin FV\ R$$

We now state the following facts about normal forms:

PROPOSITION 17.

1. *M has a $\beta\eta$-nf $\Leftrightarrow$ M has a β-nf*

2. *If M and N are distinct β-nfs then $M = N$ is not a theorem of λ (and similarly for $\lambda\eta$).*

3. *If M and N are distinct $\beta\eta$-nfs then $M \# N$.*

The use of $\beta\eta$-nfs in the last point is essential; y and $\lambda x.yx$ are distinct β-nfs but not incompatible – they are η-equivalent.

The completeness of $\lambda\eta$ is expressed by the following:

PROPOSITION 18 (Completeness).
Suppose M and N have nfs; then either:

$\lambda\eta \vdash M = N$

or

$\lambda\eta + (M = N)$ *is inconsistent*

4 REDUCTION

Normal forms can be used as canonical representatives for the convertibility equivalence classes. A more computational view results from treating normal forms as the "answers" produced from λ-term "programs". This view is justified by observing that the evaluation of the β-normal form of a term involves removing application subterms by applying the (β) rule; we have already identified this process with function application in programming languages. We will pursue this view further[7].

We illustrate the earlier discussion and motivate the following material by considering an example in a λ-calculus extended with constants. We consider the following program:

```
let
  fac 0  = 1
  fac n = n * fac(n-1)
in fac 0
```

We briefly discussed a variant of this function earlier, where we saw that it was the fixed point of a certain functional. We consider that the calculus which we are using is extended with a constant, $\mathbf{Y}$, which computes the fixed point of a given term; following the construction used in the proof of the Fixed Point Theorem, it is clear that such a constant could be defined by the following term:

$$\mathbf{Y} \equiv \lambda f.((\lambda x.f(xx))(\lambda x.f(xx)))$$

The program may be translated in the following way:

$$(\lambda f.f0)(\mathbf{Y}(\lambda fn.if(= n\ 0)1(*\ n\ (f(-\ n\ 1)))))$$

notice that the let-construct has been translated as an application term.

Consider now the normal form of the program. We can produce the normal form by repeatedly applying rule (β); in outline, we perform the

[7] Notice that in lazy functional languages such as Haskell, rather than normal forms, (weak) head normal forms are considered to be answers — we shall return to this point later.

following steps[8]:

$$
\begin{array}{rcll}
(\lambda f.f0)(\mathbf{Y}\ldots) & = & (\mathbf{Y}\ldots)0 & \\
& = & (\lambda fn.if\ldots)(\mathbf{Y}\ldots)0 & (A) \\
& = & (\lambda n.if(=n\ 0)1(*n((\mathbf{Y}\ldots)(-\ n\ 1))))0 & \\
& = & if(=0\ 0)1(*0((\mathbf{Y}\ldots)(-\ 0\ 1))) & \\
& = & if\ \mathbf{true}\ 1\ldots & \\
& = & 1 &
\end{array}
$$

Throughout this derivation we have used the convertibility relation. Convertibility is symmetrical, indeed it is an equivalence relation, but we have used it in a non-symmetrical way. We are happy to consider 1 as the answer of the above computation, the factorial of 0, but it is a little harder to see the original program as the value of the term "1". The latter view would associate an infinite set of "values" with terms such as "1". We will study some new relations between λ-terms, notably $\rightarrow_\beta$ (one step β-reduction) and $\twoheadrightarrow_\beta$ (β-reduction), the reflexive, transitive closure of $\rightarrow_\beta$. We will see that $\twoheadrightarrow_\beta$ is closely related to $=$ but is not symmetric; each $=$ in the above derivation, other than the one in step (A), could be replaced by $\rightarrow_\beta$.

In performing reduction, we are faced with a problem of strategy. For example, at line (A) there are two subterms of the form $(\lambda x.R)S$ — henceforth called β-redexes (**red**ucible **ex**pression) — as follows:

$$(\lambda fn.if\ldots)(\mathbf{Y}\ldots)0$$

and

$$(\mathbf{Y}\ldots)$$

that is, the whole term and the subterm involving the fixed point combinator. We chose to reduce the first term but consider what would happen if we consistently chose to reduce the fixed point subterm: we would never get to the answer, we would merely construct a larger and larger term! Making the "wrong" choice is not always so catastrophic, for example:

$$
\begin{array}{rcl}
(\lambda xy./(+\ x\ y)2)((\lambda z.+\ z\ 1)4)6 & \rightarrow_\beta & (\lambda y./(+((\lambda z.+\ z\ 1)4)y)2)6 \\
& \rightarrow_\beta & /(+((\lambda z.+\ z\ 1)4)6)2 \\
& \rightarrow_\beta & /(+(+\ 4\ 1)6)2
\end{array}
$$

but also:

---

[8]We have elided two steps here and used a defining property of fixed point combinators such as $\mathbf{Y}$:

$$\mathbf{Y}F = F(\mathbf{Y}F)$$

$$\begin{array}{rcl}(\lambda xy./(+\ x\ y)2)((\lambda z.+\ \ z\ 1)4)6 & \to_\beta & (\lambda xy./(+\ x\ y)2)(+\ 4\ 1)6 \\ & \to_\beta & (\lambda y./(+(+\ 4\ 1)y)2)6 \\ & \to_\beta & /(+(+\ 4\ 1)6)2\end{array}$$

and so the answer will be the same.
The above discussion should pose two questions in the reader's mind:

Question 1: Given a term and a number of reduction sequences from that term which all terminate in a normal form, is it possible that some of the sequences might terminate with different normal forms?

Question 2: Given that some choices of reduction strategy appear to be better than others in some situations (for example the bottomless pit of $(\mathbf{Y}\ldots)$) is there a best way of choosing what to do next?

The first question is closely related to the issue of *determinacy*; computationally, the question amounts to asking if we can get different answers from a program depending on how we execute it. A corollary of the Church–Rosser Theorem, which we will present below, guarantees that the answer to this question is no. The second question is less precisely formulated; the Standardisation Theorem, also presented below, addresses the question by giving a reduction order which is guaranteed to terminate with normal form if any reduction sequence does (recall $(\lambda x.xx)(\lambda x.xx)$!) but if "best" is also meant to be read as "optimal" then the question is more complicated.

4.1 Notions of Reduction

Reduction may be viewed as a special form of relation on λ-terms. Why special? Recall the discussion of the constraints on equality above; it is reasonable to place some of the same constraints on reduction. For example, if one term reduces to another, then it should do so in any context. On the other hand, bearing in mind our earlier discussion, we should not expect a reduction relation to be an equivalence relation. We introduce the following definitions:

DEFINITION 19. $R \subseteq \Lambda^2$ is a *compatible* binary relation if:

$$(M, M') \in R \Rightarrow (C[M], C[M']) \in R$$

for all $M, M' \in \Lambda$ and all contexts $C[]$ with one hole.

DEFINITION 20. $R \subseteq \Lambda^2$, is an *equality (congruence)* relation if it is a compatible equivalence relation.

DEFINITION 21. $R \subseteq \Lambda^2$, is a *reduction* relation if it is compatible, reflexive and transitive.

We call an arbitrary binary relation on Λ, a *notion of reduction.* For example, the notion of reduction that we will be particularly interested in is:

$$\beta = \{((\lambda x.M)N, M[x := N]) \mid M, N \in \Lambda\}$$

Given two notions of reduction, R_1 and R_2, we sometimes write R_1R_2 for $R_1 \cup R_2$ (notably in the case that R_1 is β and R_2 is η, we write $\beta\eta$).

The one-step reduction relation induced by some notion of reduction R, written $\rightarrow_R$, is the compatible closure of R. The closure is explicitly constructed as follows:

DEFINITION 22 (One-step R-reduction).

$$\frac{(M, N) \in R}{M \rightarrow_R N}$$
$$\frac{M \rightarrow_R N}{MZ \rightarrow_R NZ}$$
$$\frac{M \rightarrow_R N}{ZM \rightarrow_R ZN}$$
$$\frac{M \rightarrow_R N}{\lambda x.M \rightarrow_R \lambda x.N}$$

The notation "$M \rightarrow_R N$" should be read as "M R-reduces to N in one step" or "N is an R-reduct of M". We have already seen the relation $\rightarrow_\beta$, in this case we will often say that "M reduces to N in one step" or "N is a reduct of M".

The reduction relation, written $\twoheadrightarrow_R$, is the reflexive, transitive closure of the one-step reduction relation. While, as its name implies, the one-step reduction relation allows a single step of reduction, the reduction relation allows many (including zero! — allowed by reflexivity). The reflexive transitive closure is defined formally as follows:

DEFINITION 23 (R-reduction).

$$\frac{M \rightarrow_R N}{M \twoheadrightarrow_R N}$$
$$M \twoheadrightarrow_R M$$
$$\frac{M \twoheadrightarrow_R N \quad N \twoheadrightarrow_R L}{M \twoheadrightarrow_R L}$$

For the notation "$M \twoheadrightarrow_R N$", read "M R-reduces to N".

Finally, we consider R-equality (also called R-convertibility), written $=_R$. This is the equivalence relation generated by $\twoheadrightarrow_R$.

DEFINITION 24 (R-convertibility).

$$\frac{M \twoheadrightarrow_R N}{M =_R N}$$
$$\frac{M =_R N}{N =_R M}$$
$$\frac{M =_R N \quad N =_R L}{M =_R L}$$

In the case that "$M =_R N$", we say "M is R-convertible to N".

We have the following result for these relations:

PROPOSITION 25. *$\rightarrow_R$, $\twoheadrightarrow_R$ and $=_R$ are all compatible.*

Earlier, we discussed substitution and presented some results which related $=$ to the substitution operation (notably the Substitution Lemma). Similar considerations applied to $\rightarrow_R$ and $\twoheadrightarrow_R$ help us establish some differences between the two.

LEMMA 26. $N \twoheadrightarrow_R N' \Rightarrow M[x := N] \twoheadrightarrow_R M[x := N']$

The same result does not hold for $\rightarrow_R$, since the substitution may cause redexes (see below) to be duplicated. For example consider: $M \equiv xx$, $N \equiv (\lambda y.y)z$ and $N' \equiv z$, then:

$$N \rightarrow_R N'$$

but:

$$(\lambda y.y)z((\lambda y.y)z) \not\rightarrow_R zz$$

DEFINITION 27. An *R-redex* is a term M such that $(M, N) \in R$ for some term N; in this case N is called an *R-contractum* of M. A term M is called an *R-normal form (R-nf)* if it does not contain any R-redex. A term N is an R-nf of M if N is an R-nf and $M =_R N$.

If M one-step reduces to N then some sub-term of M must have been reduced; this is captured formally in the following proposition.

PROPOSITION 28. *$M \rightarrow_R N \Leftrightarrow M \equiv C[P]$, $N \equiv C[Q]$ and $(P, Q) \in R$ for some $P, Q \in \Lambda$ where $C[]$ has one hole.*

A simple corollary of this result gives us some, not unexpected, results relating reduction and normal forms:

COROLLARY 29. *Let M be an R-nf, then:*
(i) There is no N such that $M \rightarrow_R N$
(ii) $M \twoheadrightarrow_R N \Rightarrow M \equiv N$

Care should be taken with this result; it is not the case that if:

$$\forall N, M \twoheadrightarrow_R N \Rightarrow M \equiv N$$

then M is an R-nf. To see why, consider the following term with R being β:

$$M \equiv (\lambda x.xx)(\lambda x.xx)$$

We are now ready to present the Church–Rosser Theorem.

4.2 The Church–Rosser Theorem

We start by introducing the *diamond property*:

DEFINITION 30 (The Diamond Property). Let $\rhd$ be a binary relation on Λ, then $\rhd$ satisfies the diamond property, written $\rhd \models \Diamond$, if:

$$\forall M, M_1, M_2 [M \rhd M_1 \wedge M \rhd M_2 \Rightarrow \exists M_3 [M_1 \rhd M_3 \wedge M_2 \rhd M_3]]$$

If there are two diverging $\rhd$-steps from some term and $\rhd$ satisfies the diamond property, then there is always a way to converge again.

DEFINITION 31 (Church–Rosser). A notion of reduction R is said to be *Church–Rosser (CR)* if $\twoheadrightarrow_R \models \Diamond$.

We then have the following theorem:

THEOREM 32 (Church–Rosser Theorem).
Let R be CR, then:

$$M =_R N \Rightarrow \exists Z [M \twoheadrightarrow_R Z \wedge N \twoheadrightarrow_R Z]$$

The proof is by straightforward induction on the definition of $=_R$. This theorem has a useful corollary:

COROLLARY 33. *Let R be CR, then:*
(i) if N is an R-nf of M then $M \twoheadrightarrow_R N$
(ii) a term can have at most one R-nf

Thus if we can demonstrate that β is CR, we will have answered our first question. In fact the corollary tells us more; not only does it guarantee unicity of normal forms for terms, it also guarantees that if a term has a normal form then it will be possible to reduce the term to it.

To demonstrate that β is CR we must show $\twoheadrightarrow_\beta \models \Diamond$. First some notation; if $\rhd$ is some binary relation on a set X then we write $\rhd^*$ for its transitive closure and we have:

$$\rhd \models \Diamond \Rightarrow \rhd^* \models \Diamond$$

So if we could show that the reflexive closure of one-step β-reduction satisfies the diamond property, we would have finished. Alas, life is never that simple! Consider the following term:

$$(\lambda x.xx)((\lambda x.x)(\lambda x.x))$$

Then we have the following pair of divergent reductions:

$$(\lambda x.xx)((\lambda x.x)(\lambda x.x)) \rightarrow_\beta ((\lambda x.x)(\lambda x.x))((\lambda x.x)(\lambda x.x))$$

$$(\lambda x.xx)((\lambda x.x)(\lambda x.x)) \rightarrow_\beta (\lambda x.xx)(\lambda x.x)$$

But while in the second case there is then only one redex:

$$(\lambda x.xx)(\lambda x.x) \rightarrow_\beta (\lambda x.x)(\lambda x.x)$$

there is no way of converging to this term by one step in the first case. So we cannot directly apply the above result to show that β is CR. The approach that we will take, following Tait and Martin-Löf, involves introducing a new relation which is "sandwiched" by the reflexive closure of $\rightarrow_\beta$ and $\twoheadrightarrow_\beta$ and which has $\twoheadrightarrow_\beta$ as its transitive closure.

We define the relation $\twoheadrightarrow_1$. This relation is reflexive and allows multiple β-reductions in one step. We read "$M \twoheadrightarrow_1 N$" as "M grand reduces to N". The intuition is that $\twoheadrightarrow_1$ can perform multiple $\rightarrow_\beta$ steps in one big step.

DEFINITION 34 (Grand Reduction).
$\twoheadrightarrow_1$ is defined in the following way:

$$M \twoheadrightarrow_1 M$$

$$\frac{M \twoheadrightarrow_1 M'}{\lambda x.M \twoheadrightarrow_1 \lambda x.M'}$$

$$\frac{M \twoheadrightarrow_1 M' \quad N \twoheadrightarrow_1 N'}{MN \twoheadrightarrow_1 M'N'}$$

$$\frac{M \twoheadrightarrow_1 M' \quad N \twoheadrightarrow_1 N'}{(\lambda x.M)N \twoheadrightarrow_1 M'[x := N']}$$

Notice that, since $\twoheadrightarrow_1$ is reflexive, both of the divergent $\rightarrow_\beta$ steps are also $\twoheadrightarrow_1$ steps. There are two additional $\twoheadrightarrow_1$ steps, the first uses reflexivity and the second results in the term $(\lambda x.x)(\lambda x.x)$ (by using the fourth clause in the definition). Evidence that $\twoheadrightarrow_1$ is weaker than $\twoheadrightarrow_\beta$ is furnished by the fact that:

$$(\lambda x.xx)((\lambda x.x)(\lambda x.x)) \twoheadrightarrow_\beta \lambda x.x$$

but the corresponding grand reduction requires at least two steps.
The following properties of $\twoheadrightarrow_1$ can all be proved by induction on the definition of the relation:

1. $M \twoheadrightarrow_1 M', N \twoheadrightarrow_1 N' \Rightarrow M[x := N] \twoheadrightarrow_1 M'[x := N']$
2. $\lambda x.M \twoheadrightarrow_1 N \Rightarrow N \equiv \lambda x.M'$ with $M \twoheadrightarrow_1 M'$
3. $MN \twoheadrightarrow_1 L$ implies either:
   (a) $L \equiv M'N'$ with $M \twoheadrightarrow_1 M'$ and $N \twoheadrightarrow_1 N'$
   (b) or $M \equiv \lambda x.P, L \equiv P'[x := N']$ with $P \twoheadrightarrow_1 P'$ and $N \twoheadrightarrow_1 N'$
4. $\twoheadrightarrow_1 \models \Diamond$

Finally, we have the result:

THEOREM 35.
$\twoheadrightarrow_\beta$ is the transitive closure of $\twoheadrightarrow_1$

From this result and property 4 of $\twoheadrightarrow_1$ we have that β is CR.

Therefore, using the corollary to the theorem that started this section, we know that β-nfs are unique and that, if a term has a β-nf then it is possible to reduce it to that nf. This allows us to prove the consistency of the theory λ. First, we need the following:

PROPOSITION 36.
$M =_\beta N \Leftrightarrow \lambda \vdash M = N$

Consistency follows because:

$$M = N$$

is not a theorem for any two distinct normal forms (because by the Church–Rosser Theorem they would have to have a common contractum for the equality to hold).

We can also define a notion of reduction which is related to the extensional theory $\lambda\eta$:

$$\eta = \{(\lambda x.Mx, M) \mid x \notin FV(M)\}$$

We can define one-step η-reduction, η-reduction and η-convertibility in the standard way. It is then possible to address the question "Is η CR?"; however, a more interesting question is whether the derived notion $\beta\eta$ $(= \beta \cup \eta)$ is. It turns out that both η and $\beta\eta$ are CR and the interested reader is referred to Barendregt for more detail.

Newman's Lemma

An alternative strategy for proving that a notion of reduction is CR uses Newman's Lemma. First we introduce a few more definitions:

DEFINITION 37. A binary relation, $\triangleright$, on a set X satisfies the *weak diamond property* if:

$$\forall M, M_1, M_2[M \triangleright M_1 \wedge M \triangleright M_2 \Rightarrow \exists M_3[M_1 \triangleright^*_= M_3 \wedge M_2 \triangleright^*_= M_3]]$$

where $\triangleright^*_=$ is the reflexive, transitive closure of $\triangleright$.

Compare this to the diamond property; here the reductions diverge in one step but there may be many (or zero) steps for them to re-converge. The converging reduction sequences need not have the same number of steps. The reduction relation $\rightarrow_\beta$ satisfies the weak diamond property.

DEFINITION 38. R is *Weakly Church–Rosser (WCR)* if $\rightarrow_R$ satisfies the weak diamond property.

DEFINITION 39. For $M \in \Lambda$:

1. M *R-strongly normalises (R-SN(M))* if there is no infinite R-reduction starting with M.
2. M is *R-infinite (R-∞(M))* if not R-SN(M).
3. R is *Strongly Normalising (SN)* if:

$$\forall M \in \Lambda.R\text{-SN}(M)$$

Examples of (1) for β are:

$$\lambda x.x \text{ and } (\lambda x.xx)((\lambda x.x)(\lambda x.x))$$

But:

$$\beta\text{-}\infty((\lambda x.xx)(\lambda x.xx)) \text{ and } \beta\text{-}\infty((\lambda x.y)((\lambda x.xx)(\lambda x.xx)))$$

The second example is instructive since it shows that terms can be β-infinite but have normal forms. Because of the existence of these latter examples, it is clear that β is not SN and thus the following is not applicable (however it will be useful for the simply typed λ-calculus). Newman's Lemma is stated as follows:

LEMMA 40 (Newman's Lemma).

$$SN \wedge WCR \Rightarrow CR$$

Thus the alternative strategy involves showing SN and WCR separately and then inferring CR.

4.3 Delta Rules

The pure, type free λ-calculus is an extremely powerful formalism. Indeed, all computable functions are representable as λ-terms as we shall see later. Such representations use clever coding tricks. For example a possible coding of integers is equivalent to using the data type specification:

$$num = Zero \mid Succ\ num$$

so that 5 (say) is represented as $Succ(Succ(Succ(Succ(Succ\ Zero))))$ and the arithmetic operations are coded up as recursive functions, for example:

$$\begin{array}{lcl} plus(m, Zero) & = & m \\ plus(m, Succ(n)) & = & plus(Succ(m), n) \end{array}$$

An alternative to this approach is to add constants to the notation along with associated reduction rules (so-called δ-rules).

If δ is some constant, we write $\Lambda\delta$ to represent the class of terms constructed from the usual alphabet plus δ. A δ-rule is then of the form[9]:

$$\delta\vec{\mathrm{M}} \to E(\vec{\mathrm{M}})$$

An example, introduced by Church, is:

$$\begin{array}{lcl} \delta_C MN & \to & \lambda xy.x \text{ if } M, N \in \beta\delta_C\text{-}nf^0, M \equiv N \\ \delta_C MN & \to & \lambda xy.y \text{ if } M, N \in \beta\delta_C\text{-}nf^0, M \not\equiv N \end{array}$$

where $\beta\delta_C\text{-}nf^0$ are closed $\beta\delta_C$ normal forms.

Several remarks are in order. First, as we shall see later, $\lambda xy.x$ is a standard encoding for $true$ and $\lambda xy.y$ is a standard encoding for $false$. Thus δ_C is effectively a predicate which determines if two closed $\beta\delta_C$ normal forms are equivalent. It is important that the δ-rules should specify closed terms to avoid inconsistency:

$$(\lambda xy.\delta_C xy)\mathbf{II} \twoheadrightarrow \delta_C\mathbf{II} \ \to\ \lambda xy.x$$

but if δ_C can be applied to open terms then the body of the λ-expression becomes a redex and we also have:

$$(\lambda xy.\delta_C xy)\mathbf{II} \ \to\ (\lambda xyzw.w)\mathbf{II} \twoheadrightarrow \lambda zw.w$$

since $x \not\equiv y$. Now,

$$\begin{array}{lcl} \lambda xy.x & = & \lambda zw.w \\ & \Rightarrow & (\lambda xy.x)MN = (\lambda zw.w)MN \\ & \Rightarrow & M = N \end{array}$$

[9]We use the notation $E(M)$ to denote some arbitrary expression involving M.

for arbitrary M, N. For reasons that we will return to below, it is also important that δ_C operates on normal forms.

Caution is required. Even quite innocuous looking rules can disturb the Church–Rosser property. This is illustrated by the following example. We consider $\Lambda cons, head, tail$ with the rules (collectively called SP for "surjective pairing"):

$$\begin{array}{lcl} head(consM_1M_2) & \rightarrow & M_1 \\ tail(consM_1M_2) & \rightarrow & M_2 \\ cons(headM)(tailM) & \rightarrow & M \end{array}$$

Klop shows that βSP is not CR.

So how can we be sure that we will not disturb the CR property? Fortunately, there is a theorem, due to Mitschke, which gives conditions under which CR is preserved and we now present this. We start by defining what it means for two binary relations to commute:

DEFINITION 41.

Let $\rhd_1$ and $\rhd_2$ be binary relations on X. $\rhd_1$ and $\rhd_2$ *commute* if:

$$\forall x, x_1, x_2 \in X[x \rhd_1 x_1 \,\wedge\, x \rhd_2 x_2 \Rightarrow \exists x_3 \in X[x_1 \rhd_2 x_3 \,\wedge\, x_2 \rhd_1 x_3]]$$

Notice that $\rhd \models \Diamond$ if and only if $\rhd$ commutes with itself (which follows from the definition). An important (useful) lemma which makes use of this notion of commutativity is due to Hindley and Rosen:

LEMMA 42 (Hindley–Rosen Lemma).
(i) Let $\rhd_1$ and $\rhd_2$ be binary relations on X. Suppose

1. *$\rhd_1 \models \Diamond$ and $\rhd_2 \models \Diamond$*
2. *$\rhd_1$ commutes with $\rhd_2$*

then $(\rhd_1 \cup \rhd_2)^* \models \Diamond$ (where $(\rhd_1 \cup \rhd_2)^*$ is the transitive closure of the combined relation).
(ii) Let R_1 and R_2 be two notions of reduction. Suppose

1. *R_1 and R_2 are CR*
2. *$\twoheadrightarrow_{R_1}$ commutes with $\twoheadrightarrow_{R_2}$*

then $R_1 \cup R_2$ is CR.

We can now state Mitschke's theorem:

THEOREM 43.

Let δ be some constant. Let $R_1, \ldots, R_m$ be n-ary relations on $\Lambda\delta$ and let $N_1, \ldots, N_m$ be arbitrary terms in $\Lambda\delta$. Introduce the notion of reduction δ by the following rules:

$$\begin{array}{rcl} \delta\vec{M} & \to & N_1 \ \mathit{if}\ R_1(\vec{M}) \\ & & \ldots \\ \delta\vec{M} & \to & N_m \ \mathit{if}\ R_m(\vec{M}) \end{array}$$

Call this collection of rules δ_M. Then $\beta\delta_M$ is CR provided that:

1. *The R_i are disjoint*

2. *The R_i are closed under $\beta\delta_M$-reduction and substitution, that is:*

   $R_i(\vec{M}) \Rightarrow R_i(\vec{M'})$ *if* $\vec{M} \twoheadrightarrow_{\beta\delta_M} \vec{M'}$ *or* $\vec{M'}$ *is a substitution instance of* $\vec{M}$.

We refer the reader to Barendregt for details of the proof. Church's rules passes the test because of the insistence that M and N are $\beta\delta_C$ normal forms.

4.4 Residuals

In the following we will often want to trace a redex through a reduction sequence. Of course the redex, or more generally subterm, may be transformed through the sequence. For example, in the following sequence:

$$\begin{array}{l} (\lambda xy.\underline{(\lambda zw.xz)y})MN \\ \to_\beta (\lambda y.\underline{(\lambda zw.Mz)y})N \\ \to_\beta \underline{(\lambda zw.Mz)N} \\ \to_\beta \lambda w.MN \end{array}$$

the underlined redexes are clearly related though different; notice that there is no remnant of the redex in the final term (it was reduced in the preceding line). We formalise this by introducing the notion of *descendant* of a subterm; we reserve the name *residual* for the descendant of a redex. We follow Klop and Lévy by introducing these notions via a labelled variant of the λ-calculus.

Terms in the labelled λ-calculus are words over the usual alphabet plus a label set, $\mathcal{A}$ (for example $\mathcal{A}$ might be $\mathcal{Z}_{\geq 0}$ - the positive integers):

DEFINITION 44.
$\Lambda_\mathcal{A}$ is the set of labelled λ-terms defined inductively by:

1. $x^a \in \Lambda_\mathcal{A}, a \in \mathcal{A}$, x a variable

2. If $M \in \Lambda_\mathcal{A}$ and $a \in \mathcal{A}$ then $(\lambda x.M)^a \in \Lambda_\mathcal{A}$

3. If $M, N \in \Lambda_\mathcal{A}$ and $a \in \mathcal{A}$ then $(MN)^a \in \Lambda_\mathcal{A}$

For example:

$$((\lambda x.(x^1x^2)^3)^4(y^5z^6)^7)^8$$

We can develop a theory for this calculus which closely mirrors λ; rather than do that we will just define the rule (β) and associated substitution operation and leave the reader to fill in the remaining details. Since we can view labels as colours which are attached to terms and which have no effect on computation (but are preserved by reduction) the theory is similar to our earlier development. The new rule (β) is:

$$((\lambda x.A)^aB)^b = A[x := B]$$

Notice that A and B are labelled terms and their labels are preserved but the labels a and b disappear. This is reasonable since a labels the function part of the redex and b labels the redex; neither of these plays any further rôle in the reduction sequence once the redex has been reduced. The substitution operation has to respect labels:

DEFINITION 45.

$$\begin{array}{lcl} x^a[x := B] & \equiv & B \\ y^a[x := B] & \equiv & y^a, y \text{ distinct from } x \\ (MN)^a[x := B] & \equiv & (M[x := B]N[x := B])^a \\ (\lambda y.M)^a[x := B] & \equiv & (\lambda y.M[x := B])^a \end{array}$$

For example, corresponding to the unlabelled term:

$$(((\lambda x.(\lambda y.y))((\lambda x.xx)(\lambda x.xx)))((\lambda x.x)z))$$

we have the following labelled term and reduction sequence:

$$\begin{array}{l} (((\lambda x.(\lambda y.y^1)^2)^3((\lambda x.x^4x^5)^6(\lambda x.x^7x^8)^9)^{10})^{11}((\lambda x.x^{12})^{13}z^{14})^{15})^{16} \\ \to_\beta ((\lambda y.y^1)^2((\lambda x.x^{12})^{13}z^{14})^{15})^{16} \\ \to_\beta ((\lambda x.x^{12})^{13}z^{14})^{15} \\ \to_\beta z^{14} \end{array}$$

DEFINITION 46.

Let M be an unlabelled λ-term and $\mathcal{A}$ a label set. A labelling is a function, $\mathcal{I}$, mapping each subterm to a label. We call a labelling *initial* if it labels distinct subterms with distinct labels.

For a reduction Δ, we have the labelled equivalent, Δ^*:

$$\Delta^* : \mathcal{I}(M) \to^{\Delta} \mathcal{J}(N) \text{ for some labellings } \mathcal{I} \text{ and } \mathcal{J}$$

where we have used the superscript on the reduction arrow to indicate the redex that is being reduced.

DEFINITION 47.
If $\mathcal{I}(S) = \mathcal{J}(T)$ for $S \in Sub(M)$ and $T \in Sub(N)$ then T is a *descendant* of S. As already mentioned, the descendant of a redex is called a *residual* and the redex that we contract at each stage has no residuals.

4.5 Head Normal Forms

We now introduce an alternative form of normal form: *head normal form.* Head normal forms play an important rôle in the theory and they are much closer to the concept of "answer" employed in lazy functional programming languages, as we shall see.

We start with some formal definitions:

DEFINITION 48. $M \in \Lambda$ is a *head normal form (hnf)* if M is of the form:

$$\lambda x_1 \ldots x_n . x M_1 \ldots M_m \qquad n, m \geq 0$$

In this case x is called the *head variable.*
If $M \equiv \lambda x_1 \ldots x_n . (\lambda x . M_0) M_1 \ldots M_m$ where $n \geq 0, m \geq 1$ then the subterm $(\lambda x . M_0) M_1$ is called the *head redex* of M.

Some examples of head normal forms are:

- xM
- $\lambda x . x$
- $\lambda xy . x$
- $\lambda xy . x((\lambda z . z) y)$
- $\lambda y . z$

If

$$M \rightarrow^{\Delta} N$$

and Δ is the head redex of M, then we write:

$$M \rightarrow_h N$$

and we also write $\twoheadrightarrow_h$ for the many-step reduction relation.

DEFINITION 49. If A and B are two redexes in an expression M and the first occurrence of λ in A is to the left of the first occurrence of λ in B then we say that A is *to the left of* B. If A is a redex in M and it is to the left of all of the other redexes then A is the *leftmost redex.*

Notice that the head redex of a term is always the leftmost but not conversely; consider:

$$\lambda xy.x((\lambda z.z)y)$$

this term is an hnf (i.e. it has no head redex!) and so the leftmost redex is the *internal redex*[10]:

$$(\lambda z.z)y$$

Unlike normal forms, a term does not usually have a unique head normal form. For example:

$$(\lambda x.x(\mathbf{II}))z \text{ where } \mathbf{I} \equiv \lambda x.x$$

has hnf's

- $z(\mathbf{II})$
- and $z\mathbf{I}$

However, since any term has only one head redex, every term which has an hnf has a *principal head normal form* which is reached by reducing the head redex at each stage until the head normal form is reached. The principal head normal form of the example is $z(\mathbf{II})$.

Head normal forms play a crucial rôle in the Computability Theory associated with the λ-calculus. There must be some way of coding partial functions — functions which are undefined for some elements in the domain. Readers familiar with denotational semantics will have already met this problem; in domain theory, partial functions are made into total functions by adding an undefined element $\perp$ (pronounced "bottom") to the co-domain. In the λ-calculus, the solution is to use a class of terms to represent the undefined element. The first attempt at solving this problem involved equating all of the terms without normal form and then using some canonical representative as the undefined element. However, this leads to inconsistency because neither:

$$\lambda x.x\mathbf{K}\Omega \text{ where } \mathbf{K} \equiv \lambda xy.x \text{ and } \Omega \equiv (\lambda x.xx)(\lambda x.xx)$$

nor:

$$\lambda x.x\mathbf{S}\Omega \text{ where } \mathbf{S} \equiv \lambda xyz.xz(yz)$$

has an nf but it is easy to show that $\lambda + (\lambda x.x\mathbf{K}\Omega = \lambda x.x\mathbf{S}\Omega)$ is inconsistent:

$$\begin{array}{rcl} \lambda x.x\mathbf{K}\Omega = \lambda x.x\mathbf{S}\Omega & \Rightarrow & (\lambda x.x\mathbf{K}\Omega)\mathbf{K} = (\lambda x.x\mathbf{S}\Omega)\mathbf{K} \\ & \Rightarrow & \mathbf{KK}\Omega = \mathbf{KS}\Omega \\ & \Rightarrow & \mathbf{K} = \mathbf{S} \end{array}$$

[10] A redex is internal if it is not a head redex.

but we saw above that $\mathbf{K}\#\mathbf{S}$. Instead, we equate all terms which do not have a head normal form (this is a proper sub-class of the class of terms without nf); this leads to no inconsistency, a canonical representative is Ω.

Practical lazy functional programming systems even stop some way short of hnf. Most lazy systems evaluate terms to *weak head normal form.* A weak head normal form is a term of the form:

$$xM_0 \dots M_n \text{ where } n \geq 0$$

or

$$\lambda x.M$$

that is, lazy systems do not evaluate inside λs.

4.6 The Standardisation Theorem

DEFINITION 50.

A reduction sequence:

$$\sigma : M_0 \rightarrow^{\Delta_0} M_1 \rightarrow^{\Delta_1} M_2 \rightarrow^{\Delta_2} \dots$$

is a *standard reduction* if $\forall i.\forall j < i.\Delta_i$ is not a residual of a redex to the left of Δ_j relative to the given reduction from M_j to M_i.

An alternative description of a standard reduction is as follows: after reduction of each redex R, all of the λs to the left of R are marked indelibly; no redex whose first λ is marked can be further reduced.

If there is a standard reduction from some term M to some other term N then we write $M \twoheadrightarrow_s N$. Notice that any head reduction sequence is a standard reduction sequence.

We have already defined an internal redex to be any redex which is not a head redex. We write:

$$M \twoheadrightarrow_i N$$

if there is a reduction sequence:

$$M \equiv M_0 \rightarrow^{\Delta_0} M_1 \rightarrow^{\Delta_1} \dots \rightarrow^{\Delta_{n-1}} M_n \equiv N$$

such that each of the Δ_i is an internal reduction in M_i. Before we can prove the Standardisation Theorem, we must state a result which allows us to factor reductions into a sequence of head reductions followed by a sequence of internal reductions. The proof of the following result can be found in Barendregt's book.

PROPOSITION 51. $M \twoheadrightarrow N \Rightarrow \exists Z[M \twoheadrightarrow_h Z \twoheadrightarrow_i N]$

The details of the proof use some additional theory which is beyond the scope of this article; it relies on two observations:

- If $M \rightarrow_i M' \twoheadrightarrow_h N'$, then there is an equivalent reduction sequence $M \twoheadrightarrow_h N \twoheadrightarrow_i N'$.
- Any reduction sequence $M \twoheadrightarrow N$ is of the form

$$M \twoheadrightarrow_h M_1 \twoheadrightarrow_i M_2 \twoheadrightarrow_h M_3 \twoheadrightarrow_i \ldots \twoheadrightarrow_i N$$

The intuition behind the first observation is the difference between call-by-value and call-by-name: an internal redex is an argument, so if we pre-evaluate it we only need do it once, if we don't pre-evaluate it then it may be duplicated. Since any reduction is either a head reduction or an internal reduction, the second observation is straightforward.

We then have the Standardisation Theorem:

THEOREM 52 (The Standardisation Theorem).

$$M \twoheadrightarrow N \Rightarrow M \twoheadrightarrow_s N$$

Thus we are able to answer the second question posed earlier: since we know from the Corollary to the Church–Rosser Theorem that if M has a normal form N then $M \twoheadrightarrow N$, then by the Standardisation Theorem we know that a standard reduction sequence will lead to the normal form.

5 MODELS

The purpose of a model is to give a semantics for terms. The objective is to identify each term with an element of some mathematical structure, normally a set or a set with additional structure (e.g. a complete partial order); the underlying theory of the mathematical structure then becomes available as a basis for reasoning about the terms of our language and their inter-relationships.

For the type-free λ-calculus, we are unable to give a (naive) set-theoretic model. The problem is that terms serve as both functions and arguments; in particular, a term can be applied to itself — recall Ω. Consequently, a model of the type-free λ-calculus requires a structure which is isomorphic (has the same structure) as its own function space, i.e. we have to "solve" the following:

$$D \cong D \rightarrow D$$

In set theory, the only solutions are trivial (D is a singleton) which follows from consideration of the cardinalities of the sets involved. Other than the term models (see below), there were no models of the type-free λ-calculus until the late 1960s. Dana Scott realised that the isomorphism could be solved by imposing a topology on the sets and then restricting the function space to continuous functions with respect to the topology. This fundamental contribution has become known as Scott's thesis:

Scott's Thesis: All computable functions are continuous.

which has a similar status in domain theory to the Church–Turing Thesis. Scott's original work used complete lattices, his first model was called D_∞ and later he published the graph model $P\omega$. Later work in this area has tended to use sub-categories of complete partial orders[11].

A detailed treatment of any particular model, other than the term models, takes us a little far from our main theme. Instead, we will give an abstract characterisation of a model. We will introduce two classes of models:

- λ-algebras which satisfy all provable equations of the λ-calculus
- λ-models which satisfy all provable equations of the λ-calculus and the axiom of weak extensionality:

$$\forall x.(M = N) \Rightarrow \lambda x.M = \lambda x.N$$

5.1 λ-algebras

We will start with a very simple structure and successively refine it. At the very minimum, we will require a set of objects and an operation on these objects which will be used to give a semantics to application:

DEFINITION 53 (Applicative Structure).
$\mathcal{M} = (X, \bullet)$ is an *applicative structure* if $\bullet$ is a binary operation on X (i.e. $\bullet : X \times X \to X$).
$\mathcal{M}$ is said to be *extensional* if, in addition, for $a, b \in X$, one has:

$$(\forall x \in X.a \bullet x = b \bullet x) \Rightarrow a = b$$

We will usually omit the $\bullet$ and just juxtapose the "function" and its "argument" thus:

$$ax \equiv a \bullet x$$

The class of terms over an applicative structure $\mathcal{T}(\mathcal{M})$ are words over the alphabet:

$v_0, v_1, \ldots$	variables
$c_a, c_b, \ldots$	constants denoting objects in X
$(,)$	parentheses

[11] One motivation for this switch is that there is often no good computational interpretation for the Top elements which appear in the complete lattice approach.

$$\begin{array}{lcl} [\![v]\!]^{\mathcal{M}}_{\rho} & = & \rho(v) \\ [\![c_a]\!]^{\mathcal{M}}_{\rho} & = & a \\ [\![(AB)]\!]^{\mathcal{M}}_{\rho} & = & [\![A]\!]^{\mathcal{M}}_{\rho}[\![B]\!]^{\mathcal{M}}_{\rho} \end{array}$$

Figure 2. Interpretation, $[\![A]\!]^{\mathcal{M}}_{\rho}$

DEFINITION 54 (Terms).
$\mathcal{T}(\mathcal{M})$ is the least class satisfying the following:

1. $v \in \mathcal{T}(\mathcal{M})$, v a variable
2. $c_a \in \mathcal{T}(\mathcal{M})$, $a \in X$
3. if $A, B \in \mathcal{T}(\mathcal{M})$ then $(AB) \in \mathcal{T}(\mathcal{M})$

Before we can give an interpretation to terms in $\mathcal{T}(\mathcal{M})$, we need another definition. Terms can contain free variables and in order to decide what such a term denotes, we must know the "value" of the free variables. We will use an environment function to record the current bindings for the free variables:

$$\rho : variables \to X$$

An interpretation of $A \in \mathcal{T}(\mathcal{M})$ in $\mathcal{M}$ under ρ – written $[\![A]\!]^{\mathcal{M}}_{\rho}$ but we will omit ρ and the $\mathcal{M}$-superscript when they are clear from the context – is defined as shown in Figure 2
We will write:

$$\mathcal{M}, \rho \models A = B$$

read "$A = B$ is true in $\mathcal{M}$ under ρ" if:

$$[\![A]\!]^{\mathcal{M}}_{\rho} = [\![B]\!]^{\mathcal{M}}_{\rho}$$

We write $\mathcal{M} \models A = B$ and say "$A = B$ is true in $\mathcal{M}$" if

$$\mathcal{M}, \rho \models A = B \text{ for all } \rho$$

So much for the basic structure; we will now start to refine it. We introduce the following definition:

DEFINITION 55 (Combinatory Algebra).
A *combinatory algebra* is an applicative structure with two distinguished elements:

$$\mathcal{M} = (X, \bullet, k, s)$$

which satisfy:

$$kxy = x$$

and

$$sxyz = xz(yz)$$

A structure is non-trivial if its cardinality is greater than 1; a combinatory algebra is non-trivial if and only if $k \neq s$.

An arbitrary applicative structure is capable of modelling application of λ-terms but we have no obvious way of representing abstraction terms. In a combinatory algebra, it is possible to simulate abstraction and thus combinatory algebras are candidate models for the λ-calculus. We start by extending the class of terms with three distinguished constants, **K** and **S**, which denote k and s respectively and **I** which denotes $s \bullet k \bullet k$. For $A \in \mathcal{T}(\mathcal{M})$ and variable x, we define the term $\lambda^* x.A \in \mathcal{T}(\mathcal{M})$:

DEFINITION 56.

$$\begin{array}{lcl} \lambda^* x.x & \equiv & \mathbf{I} \\ \lambda^* x.P & \equiv & \mathbf{K}P, \text{if } P \text{ does not contain } x \\ \lambda^* x.PQ & \equiv & \mathbf{S}(\lambda^* x.P)(\lambda^* x.Q) \end{array}$$

It is possible to show that λ^* captures the main properties of abstraction. We extend the class of λ-terms, Λ, to $\Lambda(\mathcal{M})$ which consist of the λ-terms built from variables and constants from $\mathcal{M}$. We now define two maps which establish a relationship between $\Lambda(\mathcal{M})$ and the terms over $\mathcal{M}$:

DEFINITION 57 ($\_{}_{CL}$ and $\_{}_{\lambda}$).
$\_{}_{CL} : \Lambda(\mathcal{M}) \to \mathcal{T}(\mathcal{M})$

$$\begin{array}{lcl} x_{CL} & = & x \\ c_{CL} & = & c \\ (MN)_{CL} & = & M_{CL} N_{CL} \\ (\lambda x.M)_{CL} & = & \lambda^* x.M_{CL} \end{array}$$

$\_{}_{\lambda} : \mathcal{T}(\mathcal{M}) \to \Lambda(\mathcal{M})$

$$\begin{array}{lcl} x_{\lambda} & = & x \\ c_{\lambda} & = & c \\ \mathbf{I}_{\lambda} & = & \lambda x.x \\ \mathbf{K}_{\lambda} & = & \lambda xy.x \\ \mathbf{S}_{\lambda} & = & \lambda xyz.xz(yz) \\ (AB)_{\lambda} & = & A_{\lambda} B_{\lambda} \end{array}$$

Since we are mainly interested in λ-terms, we will abuse notation and write M when we should write M_{CL} and use the turnstile, $\models$, for equality between λ-terms:

$$\begin{array}{lcl} \mathcal{M}, \rho \models M = N & \equiv & [\![M_{CL}]\!]_\rho^{\mathcal{M}} = [\![N_{CL}]\!]_\rho^{\mathcal{M}} \\ \mathcal{M} \models M = N & \equiv & [\![M_{CL}]\!]^{\mathcal{M}} = [\![N_{CL}]\!]^{\mathcal{M}} \text{ for all } \rho \end{array}$$

DEFINITION 58 (λ-algebra).
A combinatory algebra is called a λ-*algebra* if for all $A, B \in \mathcal{T}(\mathcal{M})$:

$$\lambda \vdash A_\lambda = B_\lambda \Rightarrow \mathcal{M} \models A = B$$

Not all combinatory algebras are λ-algebras. We now give a theorem which gives a slightly more useful characterisation of λ-algebras:

THEOREM 59. *Let $\mathcal{M}$ be a combinatory algebra, then $\mathcal{M}$ is a λ-algebra iff:*
$\forall M, N \in \Lambda(\mathcal{M})$

1. $\lambda \vdash M = N \Rightarrow \mathcal{M} \models M = N$
2. $\mathcal{M} \models \mathbf{K}_{\lambda,CL} = \mathbf{K}$ *and* $\mathcal{M} \models \mathbf{S}_{\lambda,CL} = \mathbf{S}$

5.2 λ-models

Finally, we arrive at the most natural class of models: the λ-models. Given a combinatory algebra, we define:

$$\mathbf{1} = s(ki)$$

A good intuition is that $\mathbf{1}$ is a function application operator — it takes two arguments and applies the first to the second.

DEFINITION 60 (λ-model).
A λ-*model* is a λ-algebra, M, in which the following axiom, due to Meyer and Scott, holds:

$$\forall a, b, x \in M.(ax = bx) \Rightarrow \mathbf{1}a = \mathbf{1}b$$

Below, we will give an alternative characterisation of λ-models, but first we need some results about $\mathbf{1}$:

PROPOSITION 61.
Let $\mathcal{M}$ be a combinatory algebra, then in $\mathcal{M}$:

1. $\mathbf{1}ab = ab$

   If, moreover, $\mathcal{M}$ is a λ-algebra then:

2. $\mathbf{1} = \lambda xy.xy$
3. $\mathbf{1}(\lambda x.A) = \lambda x.A$ *for all* $A \in \mathcal{T}(\mathcal{M})$
4. $\mathbf{11} = \mathbf{1}$

A λ-algebra is *weakly extensional* if for $A, B \in \mathcal{T}(\mathcal{M})$:

$$\mathcal{M} \models \forall x.(A = B) \Rightarrow \lambda^* x.A = \lambda^* x.B$$

We close this section with a theorem which characterises λ-models in terms of weakly extensional λ-algebras:

THEOREM 62.
$\mathcal{M}$ is a λ-model $\Leftrightarrow$ $\mathcal{M}$ is a weakly extensional λ-algebra

5.3 Term models

As we said earlier, most models of the λ-calculus require the sets to have some kind of order-theoretic structure. There is, however, a class of models – the term models – which have a proof-theoretic structure. The basic idea is that the semantics of a term is given to be the equivalence class of the term under the convertibility relationship.

We define the equivalence class of a term M.

DEFINITION 63. $[M] \equiv \{N \in \Lambda \mid \lambda \vdash M = N\}$

As is usual, the equivalence classes partition the set of terms and we can define a quotient set:

$$\Lambda/_{\lambda} \equiv \{[M] \mid M \in \Lambda\}$$

Finally, we can define a binary operation, $\bullet$, on equivalence classes:

$$[M] \bullet [N] \equiv [MN]$$

We now have the necessary components to enable us to define a model.

DEFINITION 64 (Term Models).
The open term model for the type free λ-calculus is:

$$\mathcal{M}(\lambda) = (\Lambda/_{\lambda}, \bullet, [\lambda xy.x], [\lambda xyz.xz(yz)])$$

If it is the closed terms that are of interest, we can consider the closed term model:

$$\mathcal{M}^0(\lambda) = (\Lambda^0/_{\lambda}, \bullet, [\lambda xy.x]^0, [\lambda xyz.xz(yz)]^0)$$

We then have the following two facts:

Fact 1: $\mathcal{M}^0(\lambda)$ is a λ-algebra

Fact 2: $\mathcal{M}(\lambda)$ is a λ-model

6 COMPUTABILITY

6.1 Fixed Points

In order to study the computability aspects of the λ-calculus, we will rely extensively on the ability to construct recursive definitions. In this section we introduce the concept of a *fixed point combinator* and consider the variety of different combinators.

We start by recalling the Fixed Point Theorem:

THEOREM 65 (The Fixed Point Theorem).

$$\forall F. \exists X. X = FX$$

The proof (see earlier) inspires us to make the following definition:

DEFINITION 66 (A Fixed Point Combinator).

$$\mathbf{Y} \equiv \lambda f.(\lambda x.f(xx))(\lambda x.f(xx))$$

This is a term which, when applied to another term, is equal to the fixed point of the given term. $\mathbf{Y}$ is sometimes known as Curry's Paradoxical Combinator (consider the result of applying $\mathbf{Y}$ to a term representing logical negation). In general, any term M which satisfies the following:

$$\forall F. MF = F(MF)$$

is called a fixed point combinator – there is an infinite variety of such terms but we will only use $\mathbf{Y}$ and Θ (see below) in the following.

In the preceding paragraph we have used convertibility both in the statement of the Fixed Point Theorem and the definition of fixed point combinators. Sometimes it will be desirable to have a fixed point combinator M which satisfies the slightly stronger requirement:

$$\forall F. MF \twoheadrightarrow F(MF)$$

Notice that $\mathbf{Y}$ does not have this property but the following combinator does:

$$\Theta \equiv AA \text{ where } A \equiv \lambda xy.y(xxy)$$

since:

$$\begin{array}{rcl} \Theta F & \equiv & (\lambda xy.y(xxy))AF \\ & \to & (\lambda y.y(AAy))F \\ & \to & F(AAF) \\ & \equiv & F(\Theta F) \end{array}$$

We now introduce a result which we will make implicit use of throughout the rest of this chapter:

PROPOSITION 67. *Let* $C \equiv C(f, \vec{x})$ *be a term (with free variables* f *and* $\vec{x}$*), then:*
(i) $\exists F.\forall \vec{N}.F\vec{N} = C(F, \vec{N})$
(ii) $\exists F.\forall \vec{N}.F\vec{N} \twoheadrightarrow C(F, \vec{N})$

Proof.
In both cases, we can take $F \equiv \Theta(\lambda f\vec{x}.C(f, \vec{x}))$. Notice that we could use $\mathbf{Y}$ instead for (i). ■

EXAMPLE 68. As an example suppose that:

$$C \equiv fyxf \equiv C(f, x, y)$$

then (i) guarantees the existence of a term F such that:

$$Fxy = FyxF$$

Just take $F \equiv \Theta(\lambda fxy.fyxf)$ then:

$$\begin{array}{rcl} Fxy & \equiv & \Theta(\lambda fxy.fyxf)xy \\ & = & (\lambda fxy.fyxf)(\Theta(\lambda fxy.fyxf))xy \\ & \equiv & (\lambda fxy.fyxf)Fxy \\ & = & FyxF \end{array}$$

EXAMPLE 69. A more familiar example is:

$$C \equiv \mathbf{if}\ n = 0\ \mathbf{then}\ 1\ \mathbf{else}\ n \times f(n-1) \equiv C(f, n)$$

and (i) guarantees the existence of a term, F, which behaves like a factorial function[12], i.e.:

$$Fn = \mathbf{if}\ n = 0\ \mathbf{then}\ 1\ \mathbf{else}\ n \times F(n-1)$$

and we just take:

$$F \equiv \mathbf{Y}(\lambda fn.\mathbf{if}\ n = 0\ \mathbf{then}\ 1\ \mathbf{else}\ n \times f(n-1))$$

Finally we recall the definition of the term Ω:

$$\Omega \equiv \omega\omega \text{ where } \omega \equiv \lambda x.xx$$

and just note that[13]:

$$\Omega = \mathbf{YI}$$

[12]Of course we have deviated somewhat from the standard syntax for terms but hopefully the message in this example is clear.
[13]Readers familiar with domain theory should consider what the fixed point of the identity function is. Ω is playing the same role as $\perp$.

6.2 Numeral Systems

In the next subsection we show the equivalence of the λ-calculus and Recursive Function Theory. To do this, we will need to define λ-terms which encode numerals, booleans, conditionals and various other constructs; we shall consider various approaches to this problem in this section.

We start with boolean values. We define true and false by terms $\mathbf{T}$ and $\mathbf{F}$:

DEFINITION 70 (True and False).

$$\begin{array}{rcl}\mathbf{T} & \equiv & \lambda xy.x \equiv \mathbf{K} \\ \mathbf{F} & \equiv & \lambda xy.y \equiv \mathbf{KI}\end{array}$$

These choices are motivated by the simple definition of the conditional function:

$$\mathbf{if} \equiv \lambda pca.pca$$

since:

$$\mathbf{if}\ \mathbf{T}MN \twoheadrightarrow M \text{ and } \mathbf{if}\ \mathbf{F}MN \twoheadrightarrow N$$

There are also simple representations for the standard boolean operations, for example **and**[14]:

$$\mathbf{and} \equiv \lambda xy.xy\mathbf{F}$$

We will also need to manipulate pairs of terms or, more generally, tuples.

DEFINITION 71 (Pairs).
We define the pairing operation as a distfix operator, $[\_,\_]$:

$$[M, N] \equiv \lambda z.zMN$$

The first and second projection functions on a pair are defined as

$$(\lambda p.p\mathbf{T}) \text{ and } (\lambda p.p\mathbf{F})$$

respectively.

These definitions are sensible since, for example, if $M \equiv [P, Q]$ then:

$$\begin{array}{rcl}M\mathbf{T} & \equiv & (\lambda z.zPQ)\mathbf{T} \\ & \rightarrow & \mathbf{T}PQ \\ & \twoheadrightarrow & P\end{array}$$

[14]This encoding uses a trick that is often used in the code generators of compilers, which is to encode the logical operations as conditional expressions. For example **and** x y is equivalent to:

if x **then** y **else** false

Ordered n-tuples can now be defined using pairing:

$$\begin{array}{rcl} [M] & \equiv & M \\ [M_0, \ldots, M_{n+1}] & \equiv & [M_0, [M_1, \ldots, [M_n, M_{n+1}] \ldots]] \end{array}$$

The generalisation of the projection functions are defined by the following terms; $\pi_{i,n}$ selects the i-th element from an $n+1$ element tuple, $0 \leq i < n$:

$$\begin{array}{rcl} \pi_{i,n} & \equiv & \lambda x.x\mathbf{F}^{*i}\mathbf{T} \equiv \lambda x.x \ldots (\text{i occurrences of } \mathbf{F}) \ldots \mathbf{T} \\ \pi_{n,n} & \equiv & \lambda x.x\mathbf{F}^{*n} \end{array}$$

An alternative approach to defining tuples is slightly more direct:

$$< M_0, \ldots, M_n > \equiv \lambda z.zM_0 \ldots M_n$$

and then we define the projection functions as follows:

$$P_{i,n} \equiv \lambda x.xU_{i,n}$$

where

$$U_{i,n} \equiv \lambda x_0 \ldots x_n.x_i$$

Before we introduce our first numeral system, we need one further combinator, composition, which is written as an infix operator:

$$M \circ N \equiv \lambda x.M(Nx)$$

We now define the numerals as the following terms:

DEFINITION 72 (Standard Numerals).

$$\begin{array}{rcl} \ulcorner 0 \urcorner & \equiv & \mathbf{I} \\ \ulcorner n+1 \urcorner & \equiv & [\mathbf{F}, \ulcorner n \urcorner] \end{array}$$

So for example:

$$\ulcorner 3 \urcorner \equiv [\mathbf{F}, [\mathbf{F}, [\mathbf{F}, \mathbf{I}]]]$$

This way of constructing numerals is reminiscent of the following type construction:

$$num = Zero \mid Succ\, num$$

in which 3 would be represented as:

$$Succ(Succ(Succ\, Zero))$$

This motivates the definition of a successor function $\mathbf{S}^+$:

$$\mathbf{S}^+ \equiv \lambda x.[\mathbf{F}, x]$$

The predecessor function, which decrements the numeral by 1, is just the second projection function:

$$\mathbf{P}^- \equiv \lambda x.x\mathbf{F}$$

Notice that:

$$\mathbf{P}^-(\ulcorner 0 \urcorner) \equiv \mathbf{P}^-\mathbf{I} \to \mathbf{IF} \to \mathbf{F}$$

We also define a unary predicate, **Zero**, which returns **T** if its argument is $\ulcorner 0 \urcorner$ and **F** otherwise:

$$\mathbf{Zero} \equiv \lambda x.x\mathbf{T}$$

since:

$$\begin{array}{lcl} \mathbf{IT} & = & \mathbf{T} \\ [\mathbf{F}, n]\mathbf{T} & = & \mathbf{F} \end{array}$$

Given this encoding and the two functions and the predicate, we can define more sophisticated functions such as addition:

$$+xy = \mathbf{if}(\mathbf{Zero}\ x)y(+(\mathbf{P}^-x)(\mathbf{S}^+y))$$

(use Proposition 67).

This encoding for numerals is by no means the only possibility. Before introducing another encoding we make some definitions:

DEFINITION 73.

A *numeral system* is a sequence:

$$\mathbf{d} = d_0, d_1, \ldots$$

consisting of closed terms such that there are λ-terms, S_d^+ and $Zero_d$ such that:

$$\begin{array}{lcl} S_d^+\ d_n & = & d_{n+1} \\ Zero_d\ d_0 & = & \mathbf{T} \\ Zero_d\ d_{n+1} & = & \mathbf{F} \end{array}$$

for all numbers n, i.e. we have codes for all numerals, the successor function and a test for zero.

DEFINITION 74. $\mathbf{d}$ is a *normal* numeral system if each d_n has a normal form.

DEFINITION 75. $\mathbf{s} = \ulcorner 0 \urcorner, \ulcorner 1 \urcorner, \ldots$ with successor function $\mathbf{S}^+$ is called the *standard* numeral system.

It is clear that the numerals in the standard numeral system are all distinct normal forms; thus the standard numeral system is a normal system.

The system $\mathbf{d}$ is determined by d_0 and S_d^+, so we often write:

$$\mathbf{d} = (d_0, S^+)$$

A foretaste of the next section is given by the following definition:

DEFINITION 76.

Let **d** be a numeral system, a numeric function:

$$\phi : N^p \to N$$

(where N is the set of natural numbers) is *λ-definable with respect to* **d** if:

$$\exists F.\forall n_1, \ldots, n_p \in N.F d_{n_1} \ldots d_{n_p} = d_{\phi(n_1,\ldots,n_p)}$$

We say that **d** is *adequate* if and only if all recursive functions are λ-definable with respect to **d**. Alternatively, **d** is adequate if and only if we can define a predecessor function for **d**.

An alternative encoding of the numerals is due to Church:

DEFINITION 77 (Church Numerals). $\mathbf{c} = c_0, c_1, \ldots$

$$c_n = \lambda f x.f^n(x)$$

The successor function is defined by:

$$S^+_{\mathbf{c}} c_n \equiv \lambda abc.b(abc)$$

We can define translation functions between the standard and Church numerals, H and H^{-1}, such that:

$$H\ulcorner n \urcorner = c_n \text{ and } H^{-1} c_n = \ulcorner n \urcorner$$

These functions are defined in the following way:

$$\begin{array}{lcl} Hx & = & \mathbf{if}\ (\mathbf{Zero}\ x)\ c_0\ S^+_{\mathbf{c}}(H(\mathbf{P}^- x)) \\ H^{-1} c_n & = & c_n\ \mathbf{S}^+(\ulcorner 0 \urcorner) \end{array}$$

Given these, we can define a test-for-zero:

$$Zero_{\mathbf{c}} \equiv \mathbf{Zero} \circ H^{-1}$$

The Church numeral system is also adequate, since we can define a predecessor function:

$$P^-_{\mathbf{c}} \equiv H \circ \mathbf{P}^- \circ H^{-1}$$

The Church numerals are of interest because we can define some of the more powerful arithmetic functions without recursion; for example $x \circ y$ gives the result of multiplying the Church numeral x by the numeral y.

6.3 λ-definability

We can specialise Definition 76 to the standard numeral system. In this case we talk about a numeric function being λ-definable (without specifying a numeral system). Since standard numerals are normal forms, in particular $\ulcorner\phi(n_1,\ldots,n_p)\urcorner$ is a normal form, we also have, by the Church–Rosser theorem, that:

$$F\ulcorner n_1\urcorner\ldots\ulcorner n_p\urcorner \twoheadrightarrow \ulcorner\phi(n_1,\ldots,n_p)\urcorner$$

Our definition implicitly assumes that the given numeric function is total, i.e. defined on its whole domain. The results can be extended to partial functions but we will mainly consider total functions in this section; there is a brief discussion of partial functions at the end of the section. We start by defining the class of total recursive functions and then proceed to demonstrate that the functions in this class are all λ-definable.

DEFINITION 78 (Initial Functions).
We define the following numeric functions to be the *initial* functions:

$$\begin{array}{lcll} U_{i,p}(n_0,\ldots,n_p) & = & n_i & 0\leq i\leq p \\ S^+(n) & = & n+1 & \\ Z(n) & = & 0 & \end{array}$$

i.e. a family of selector functions, a successor function and a constant zero function.

If $P(n)$ is a numeric relation, we use the notation:

$$\mu m[P(m)]$$

to denote the least number m for which $P(m)$ holds; or to denote undefined if there is no such m.

Given a class of numeric functions, A, we consider the following closure operators on the class:

DEFINITION 79.

- A is *closed under composition* if for all ϕ defined by:

  $$\phi(\vec{n}) = H(G_1(\vec{n}),\ldots,G_m(\vec{n}))$$

  with $H, G_1,\ldots,G_m \in A$, one has $\phi \in A$.

- A is *closed under primitive recursion* if for all ϕ defined by:

  $$\begin{array}{lcl} \phi(0,\vec{n}) & = & H(\vec{n}) \\ \phi(k+1,\vec{n}) & = & G(\phi(k,\vec{n}),k,\vec{n}) \end{array}$$

  with $H, G \in A$, one has $\phi \in A$.

- A is *closed under minimalisation* if for all ϕ defined by:

  $$\phi(\vec{n}) = \mu m[H(\vec{n}, m) = 0]$$

  with $H \in A$, such that[15]:

  $$\forall \vec{n}.\exists m.H(\vec{n}, m) = 0$$

  one has $\phi \in A$.

Notice that the primitive recursion construction is similar to the for-loop construction found in Algol-like languages and the DO-loop of FORTRAN; it provides iteration for a predetermined number of steps. It is possible to define most of the basic arithmetic functions using primitive recursion, for example:

$$\begin{array}{lcl} plus(0, y) & = & id(y) \\ plus(k+1, y) & = & F(plus(k, y), k, y) \\ & & \text{where } F(x, y, z) = S^{+}(U_{0,2}(x, y, z)) \end{array}$$

where *id* is the identity function. In contrast, the minimalisation construct corresponds to the more general form of iteration represented by `while`...`do`...and `repeat`...`until`...loops in Algol-like languages.

The class of *recursive* functions may now be defined formally as the least class of numeric functions which contains all of the initial functions and is closed under composition, primitive recursion and minimalisation.

We will now demonstrate that the initial functions are λ-definable and that the class of λ-definable functions is appropriately closed. First, we define:

$$\begin{array}{lcl} U_{i,p} & \equiv & \lambda x_0 \ldots x_p.x_i \\ S^{+} & \equiv & \lambda x.[\mathbf{F}, x] \\ Z & \equiv & \lambda x.\ulcorner 0 \urcorner \end{array}$$

Now suppose that $H, G_1, \ldots, G_m$ are λ-defined by $S, T_1, \ldots, T_m$, then:

$$\phi(\vec{n}) = H(G_1(\vec{n}), \ldots, G_m(\vec{n}))$$

is λ-defined by:

$$F \equiv \lambda \vec{x}.S(T_1\vec{x}) \ldots (T_m\vec{x})$$

If ϕ is defined by:

$$\begin{array}{lcl} \phi(0, \vec{n}) & = & H(\vec{n}) \\ \phi(k+1, \vec{n}) & = & G(\phi(k, \vec{n}), k, \vec{n}) \end{array}$$

[15]This condition ensures that ϕ is total.

with H and G λ-defined by S and T respectively, then ϕ is λ-defined by:

$$F \equiv \mathbf{Y}\lambda f x\vec{y}.(Zero\ x)\ (S\vec{y})\ (T(f(\mathbf{P}^{-}x)\vec{y})(\mathbf{P}^{-}x)\vec{y})$$

In order to define minimalisation, we first define a function which, given a predicate λ-defined by P, determines the least numeral which satisfies P. We start by defining:

$$H_P \equiv \Theta(\lambda hz.(Pz)z(h(S^{+}z)))$$

which just iterates from a given numeral, z, until (Pz) is true and returns z. The required function, written μP, is defined thus:

$$\mu P \equiv H_P \ulcorner 0 \urcorner$$

Then suppose that ϕ is defined by:

$$\phi(\vec{n}) = \mu m[H(\vec{n}, m) = 0]$$

where H is λ-defined by S; then ϕ is λ-defined by:

$$F \equiv \lambda\vec{x}.\mu[\lambda y.Zero(S\vec{x}y)]$$

From the preceding paragraphs, we have that the initial functions are λ-definable and that the three function-forming operations can be encoded in the λ-calculus. Consequently, we can infer that all (total) recursive functions are λ-definable. We also have the following result:

THEOREM 80.
If ϕ is λ-defined by F, then $\forall\vec{n}, m \in N$:

$$\phi(\vec{n}) = m \Leftrightarrow F\ \ulcorner \vec{n} \urcorner = \ulcorner m \urcorner$$

Putting these two results together, we get the following theorem (due to Kleene):

THEOREM 81.
The λ-definable numeric functions are exactly the recursive functions.

Finally, the definition can be extended to partial functions in the following way:

DEFINITION 82. A partial numeric function, ϕ, with p arguments is λ-definable if for some $F \in \Lambda$:

$$\begin{array}{lll} \forall\vec{n} \in N^p. & & \\ F\ \ulcorner \vec{n} \urcorner & = & \ulcorner \phi(\vec{n}) \urcorner \text{ if } \phi(\vec{n}) \text{ converges (i.e. is defined)} \\ F\ \ulcorner \vec{n} \urcorner & \multicolumn{2}{l}{\text{without hnf otherwise}} \end{array}$$

where $\vec{n} \equiv n_1, \ldots, n_p$

In this section we have characterised the class of functions which are λ-definable. In general, the link between λ-definability and Recursive Function Theory is:

$$\phi \text{ is } \lambda\text{-definable} \Leftrightarrow \phi \text{ is partial recursive}$$

Given another result from Computability Theory:

$$\phi \text{ is partial recursive} \Leftrightarrow \phi \text{ is Turing Computable}$$

we see that λ-definability, according to the Church–Turing thesis, can be claimed to capture the notion of *effective calculability.*

6.4 Decidability

One of the fundamental theorems of Mathematical Logic is Gödel's Incompleteness Theorem; the details of the theorem are tangential to this article but the proof of the theorem uses a coding technique which gives an effective way of associating a unique integer, the Gödel number, with each sentence in some theory. Translating this result to the λ-calculus, we have that there is an algorithmic injective map $\# : \Lambda \to N$ such that $\#M$ is the Gödel number of M. Using this notion, we can state the Second Fixed Point Theorem:

THEOREM 83 (The Second Fixed Point Theorem).

$$\forall F.\exists X.F\ulcorner\#X\urcorner = X$$

Proof.
Define:

$$\begin{array}{lcl} \mathbf{Ap}\ulcorner\#M\urcorner(\ulcorner\#N\urcorner) & = & \ulcorner\#(MN)\urcorner \\ \mathbf{Num}\ulcorner\#n\urcorner & = & \ulcorner\#(\ulcorner\#n\urcorner)\urcorner \end{array}$$

Now take $W \equiv \lambda x.F(\mathbf{Ap}\ x(\mathbf{Num}\ x))$ and $X \equiv W\ulcorner\#W\urcorner$; then:

$$\begin{array}{lcl} X & \to & F(\mathbf{Ap}\ulcorner\#W\urcorner(\mathbf{Num}\ulcorner\#W\urcorner)) \\ & = & F(\mathbf{Ap}\ulcorner\#W\urcorner(\ulcorner\#\ulcorner\#W\urcorner\urcorner)) \\ & = & F(\ulcorner\#(W\ulcorner\#W\urcorner)\urcorner) \\ & \equiv & F\ulcorner\#X\urcorner \text{ as required} \end{array}$$

■

Notice how this construction parallels the proof of the Fixed Point Theorem. Its importance for us is that it allows us to prove Scott's Theorem (the analogue of Rice's Theorem) and thereby answer some important questions about decidability in the λ-calculus.

In the following we assume that A and B are subsets of λ-terms:

DEFINITION 84. A is *non-trivial* if $A \neq \emptyset$ and $A \neq \Lambda$.

DEFINITION 85. A is *closed under equality* if:

$$\forall M, N \in \Lambda[M \in A \wedge M = N \Rightarrow N \in A]$$

DEFINITION 86. A and B are *recursively separable* iff there is a recursive set[16] C such that:

$$(A \subseteq C) \wedge (B \cap C = \emptyset)$$

Scott's Theorem may be stated thus:

THEOREM 87 (Scott's Theorem).

1. *Let A and B, subsets of Λ, be non-empty sets closed under equality. Then A and B are not recursively separable.*

2. *Let A, a subset of Λ, be a non-trivial set closed under equality. Then A is not recursive.*

We can now show the undecidability of the question as to whether an arbitrary term has a normal form — this is equivalent, in some senses, to the Halting Problem for Turing Machines. The theorem is formally stated:

THEOREM 88.
$\{M \mid M$ *has a nf*$\}$ *is an recursively enumerable*[17] *set which is not recursive.*

We can also show the undecidability of λ. First we define the notion of *essential undecidability*:

DEFINITION 89. A theory $\mathcal{T}$ is *essentially undecidable* iff $\mathcal{T}$ is consistent and has no consistent recursive extension.

The theorem is then:

THEOREM 90.
λ *is essentially undecidable*

[16] By the term "recursive set" we mean a set whose membership predicate is recursive; i.e. there is a Turing machine which for any potential element either halts with an indication that the element is a member or halts with a contrary indication.

[17] A set is *recursively enumerable* if we can construct a Turing machine which, given a potential element, will stop with the answer YES if the element is in the set but may not halt otherwise.

Proof.
Let $\mathcal{T}$ be a consistent extension of λ, then let $X = \{M \mid \mathcal{T} \vdash M = \mathbf{I}\}$.
X is not empty because surely $\mathcal{T} \vdash \mathbf{I} = \mathbf{I}$!
$X \neq \Lambda$ because $\mathcal{T}$ is consistent.
X is clearly closed under equality.
Thus, by Scott's Theorem (ii), X is not recursive and thus $\mathcal{T}$ is not recursive.
■

7 TYPED CALCULI

7.1 Typed λ-calculus

We start our study of typed calculi with the simply typed λ-calculus; this calculus has a strong typing discipline similar to that adopted in many typed imperative and object-oriented languages — each term has a single (monomorphic) type associated with it. The simply typed λ-calculus is in many ways simpler than the (type-free) λ-calculus; for example self-application is outlawed and thus all terms are strongly normalising and there are no fixed point combinators. Once again, in introducing a new calculus, we should address all of the issues that we have considered for the λ-calculus (reduction, models, computability, etc...) but instead we will just present the highlights.

There are two approaches that can be taken in defining a typed calculus. The first, originated by Curry, is called *implicit* typing; the terms are the same as the type-free calculus and each term has a set of possible types assigned to it. The second approach, originated by Church, is called *explicit* typing; terms are annotated with type information which uniquely determines a type for the term. In the following, we will follow Church's approach.

Since terms will have types associated with them, we start by considering the syntax of types:

DEFINITION 91 (Types).
The set of types, Typ, is the least set such that:

1. $0 \in Typ$
2. if $\sigma, \tau \in Typ$ then $(\sigma \rightarrow \tau) \in Typ$

The type 0 is a *ground* type. Notice that we only have a single ground type; later we will see that it plays the role of a type variable. In a more realistic language, we might differentiate between type constants and variables; for example in a programming language context, the type constants are the "built-in" types such as integers, booleans and characters. However, since we are considering a pure calculus it is sufficient to restrict ourselves to a single ground type. Types of the form $(\sigma \to \tau)$ correspond to a function type; a function of this type takes arguments of type σ and returns a result of type τ. Examples of types are:

$$0 \qquad (0 \to 0) \qquad ((0 \to 0) \to (0 \to 0))$$

If we adopt the convention that $\to$ associates to the right[18], we can omit the majority of the parentheses:

$$0 \qquad 0 \to 0 \qquad (0 \to 0) \to 0 \to 0$$

Terms in the typed λ-calculus are words over the alphabet:

$v_0^\sigma, v_1^\sigma, \ldots$	variables, a distinct set for each $\sigma \in Typ$
λ	
$(,)$	parentheses

The class of typed λ-terms is written Λ^τ; when we want to talk about the class of terms of some specific type, σ, we write Λ_σ.

DEFINITION 92 (Typed Terms).
The class Λ^τ is the class:

$$\bigcup\{\Lambda_\sigma \mid \sigma \in Typ\}$$

and the Λ_σ are such that:

1. $v_i^\sigma \in \Lambda_\sigma$
2. $M \in \Lambda_{\sigma\to\tau}, N \in \Lambda_\sigma \Rightarrow (MN) \in \Lambda_\tau$
3. $M \in \Lambda_\tau, x \in \Lambda_\sigma \Rightarrow (\lambda x.M) \in \Lambda_{\sigma\to\tau}$

Free/bound variables, closed terms and substitution are defined in the obvious way (by analogy to the type-free calculus). Care must be taken to respect the types; for example:

$$FV(\lambda v^0.v^{0\to0}) = \{v^{0\to0}\}$$

The theories λ^τ and $\lambda\eta^\tau$[19] are defined in the same way as the corresponding type-free theories but the types of terms have to make sense, for example:

$$(\lambda x^\sigma.M)N = M[x^\sigma := N] \text{ if } N \in \Lambda_\sigma$$

[18] A moment's thought should convince the reader that this convention is consistent with the left associativity of application.

[19] In future we will write $\lambda(\eta)^\tau$ to stand for either of these theories.

and formulae are of the form:

$$M = N \text{ with } M, N \in \Lambda_\sigma \text{ for arbitrary type } \sigma$$

Notions of reduction in the typed λ-calculus are the obvious analogues of the notions that we introduced in the type-free case:

$$\beta = \{(((\lambda x^\sigma.M)N), M[x^\sigma := N]) \mid M \in \Lambda_\tau, N \in \Lambda_\sigma \text{ for some } \sigma, \tau \in Typ\}$$

$$\eta = \{((\lambda x^\sigma.Mx^\sigma), M) \mid M \in \Lambda_{\sigma\to\tau} \text{ for some } \sigma, \tau \in Typ, x^\sigma \notin (FV M)\}$$

By analogy with the type-free case we have that $\beta(\eta)$ is CR.

Strong Normalisation

An essential difference between the type-free and the typed calculus is that, in the latter case, $\beta(\eta)$ is strongly normalising (SN), i.e. $\beta\eta$-SN. As a consequence of the strong normalisation result, all typed terms have normal forms; moreover, provable equality in $\lambda(\eta)^\tau$ is decidable:

PROPOSITION 93. $\lambda(\eta)^\tau \vdash M = N$ *implies* M *and* N *have the same* $\beta(\eta)$-*nfs.*
The nfs can be found effectively by SN.

It should be fairly obvious that many type-free terms can be given a type (or many types!). For example:

$$\lambda x.x$$

can be typed as:

$$\lambda x^\sigma.x^\sigma \in \Lambda_{\sigma\to\sigma} \text{ for all } \sigma \in Typ$$

that is: "$\sigma \to \sigma$ is a possible type for $\lambda x.x \in \Lambda$". However, there are many terms that can not be assigned a type; given our earlier comments and the structure of the typed terms, it should be clear that any term involving self-application falls into this category, for example:

- In order to assign a type to $\lambda x.xx$, we must assign a type to xx.
- In order to assign a type to xx, x must have type $\alpha \to \beta$ <u>and</u> type α.

Suppose $M \in \Lambda_\sigma$, we write $\mid M \mid$ ($\in \Lambda$) for the term produced by erasing all of the type symbols in M; clearly, $\mid M \mid$ is typable and a possible type is σ. If σ is a type then σ^* is an instance of σ if it results from σ by replacing some of the 0's in σ by some other type:

EXAMPLE 94. Some instances of 0:

$$0 \to 0, \qquad 0 \to 0 \to 0, \qquad ((0 \to 0 \to 0) \to (0 \to 0) \to 0 \to 0)$$

Some instances of $0 \to 0$:

$$(0 \to 0) \to 0 \to 0, \qquad (0 \to 0 \to 0) \to 0 \to 0 \to 0$$

Two important results concerning these issues were first discovered by Roger Hindley; we state them without proof:

PROPOSITION 95.

1. *The set of typable λ-terms is recursive; i.e. there is an algorithm which will decide whether a given term is typable or not.*

2. *If $M \in \Lambda$ is typable then one can find a unique $\sigma \in Typ$ such that every possible type for M is an instance of σ; σ is called the* principal type scheme *for M.*

Since we cannot have fixed point combinators in the typed λ-calculus, the reader may have wondered about the impact this has on computability. We can define a notion of λ^τ-definability analogously to λ-definability but there are some problems. The first problem is that we cannot use the standard numerals:

$$\begin{array}{lcl} \ulcorner 0 \urcorner & \equiv & \mathbf{I} \\ \ulcorner n+1 \urcorner & \equiv & [\mathbf{F}, \ulcorner n \urcorner] \end{array}$$

since the numerals all have different types, for example:

$$\begin{array}{lll} \ulcorner 0 \urcorner \equiv \mathbf{I} & \text{has type} & 0 \to 0 \\ \ulcorner 1 \urcorner \equiv \lambda z.z\mathbf{FI} & \text{has type} & ((0 \to 0 \to 0) \to (0 \to 0) \to 0) \to 0 \end{array}$$

As a consequence, the successor function (for example) is untypable! However, the Church numerals all have the same type:

$$c_n \equiv \lambda f x.f^n x$$

$$c_n \in \Lambda_{(0 \to 0) \to 0 \to 0}$$

DEFINITION 96. The *extended polynomials* are the least class of numeric functions containing:

1. Projections: $U_{i,n}$

2. Constant functions

3. The sg function : $sg\ 0 = 0,\ sg\ (n+1) = 1$

and is closed under addition and multiplication.

It turns out that it is exactly this class of functions which is λ^τ-definable on the Church numerals; the interested reader is referred to the literature for the proof. Using constant functions, addition and multiplication it is possible to construct functions which have polynomial expressions as bodies. The adjective "extended" is used in the definition to indicate that we can encode conditional functions using the sg function (with multiplication).

7.2 The Polymorphic λ-calculus

This calculus was invented independently by Girard and Reynolds. Just as the λ-calculus and functional programming have been sloganised by "functions as first class citizens" (Stoy), the 2nd-order λ-calculus can be sloganised by "Types as first class citizens"; types can be abstracted just as normal values:

EXAMPLE 97 (A polymorphic identity function:).

$$M \equiv \Lambda t.\lambda x \in t.x$$

we can then specialise this term to a particular type by application:

$$Mint \text{ or } M[int]$$

Type schemes in this calculus are constructed in the following way:

$$\sigma ::= \alpha \mid \iota \mid \sigma_1 \to \sigma_2 \mid \forall\alpha.\sigma$$

where α is a type variable and ι is a type constant. The last component is the type scheme associated with Λ-abstractions.

DEFINITION 98. The terms of the 2nd-order polymorphic λ-calculus, Λ_2, are the least class such that:

1. Every variable and constant is in Λ_2.

2. $M, N \in \Lambda_2 \Rightarrow (MN) \in \Lambda_2$.

3. $M \in \Lambda_2$, x a variable, σ a type scheme $\Rightarrow (\lambda x \in \sigma.M) \in \Lambda_2$.

4. $M \in \Lambda_2$, σ a type scheme $\Rightarrow (M\sigma) \in \Lambda_2$.

5. $M \in \Lambda_2$, α a type variable $\Rightarrow (\Lambda\alpha.M) \in \Lambda_2$.

$$A \vdash x : \sigma \qquad\qquad (x : \sigma \text{ in } A)$$

$$\frac{A_x \cup \{x : \sigma\} \vdash M : \tau}{A \vdash (\lambda x \in \sigma.M) : \sigma \to \tau}$$

$$\frac{A \vdash M : \sigma \to \tau \qquad A \vdash N : \sigma}{A \vdash (MN) : \tau}$$

$$\frac{A \vdash M : \sigma}{A \vdash (\Lambda t.M) : \forall t.\sigma} \qquad t \notin FV(A)$$

$$\frac{A \vdash M : \forall t.\sigma}{A \vdash (M\tau) : [\tau/t]\sigma}$$

where A_x is the same as A except any assumption about x has been removed.

Figure 3. Type inference for the 2nd-order polymorphic λ-calculus.

Substitution and α-congruence are defined in the obvious way. We have two β-conversion axioms:

$$\begin{array}{lrcl} (\beta^1) & (\lambda x \in \sigma.M)N & = & M[x := N] \\ (\beta^2) & (\Lambda t.M)\sigma & = & M[t := \sigma] \end{array}$$

Of course it is also possible to define η-conversion. Some basic facts concerning $\beta\eta$-reduction are:

- $\beta\eta$ is CR
- Every Λ_2 term has a $\beta\eta$-nf
- $\beta\eta$ is SN

We now present a formal system for type inference in the 2nd-order polymorphic λ-calculus. Basic judgements have the following form:

$$A \vdash e : \sigma$$

where A is a list of assumptions of the form $x : \sigma$ assigning types to variables. The axioms and rules are shown in Figure 3.

For example, we have:

$$\cfrac{\cfrac{x : \alpha \vdash x : \alpha}{\vdash (\lambda x \in \alpha.x) : \alpha \to \alpha}}{\vdash (\Lambda\alpha\lambda x \in \alpha.x) : (\forall\alpha.\alpha \to \alpha)}$$

Reynolds has used the second-order polymorphic λ-calculus to model various programming language concepts such as type definitions, abstract data types and polymorphism.

The style of polymorphism found in most functional programming languages (discussed below) is a restricted version of that discussed above. In particular, the syntax of types in Λ_2 allows arbitrary nesting of quantifiers. For example, the following is a valid type in Λ_2:

$$\forall\alpha.(\forall\beta.\alpha \to \beta) \to \alpha \to \alpha$$

The type schemes assigned to terms in functional programming systems are usually *shallow*; the quantifiers are usually omitted and are implicitly at the outermost level. Consequently, the scope of all quantifiers is the whole scheme (to the right of the quantifier).

7.3 Polymorphic Type Inference

A simple example of a polymorphic function in a functional programming language is:

$$map : (* \to **) \to [*] \to [**]$$

$$\begin{array}{lcl} map\ f\ [] & = & [] \\ map\ f\ (a : x) & = & (f\ a) : (map\ f\ x) \end{array}$$

The symbols $*$ and $**$ are used as type variables and [ ] is the list type constructor. One of the first programming languages to allow polymorphic functions was *ML* and in this section we introduce an algorithm, due to Milner, which given an untyped function will either find a polymorphic type for it or indicate that it is untypable.

We will consider the following language *Exp* of expressions:

$$e ::= x \mid e\ e' \mid \lambda x.e \mid let\ x\ =\ e\ in\ e'$$

Types are constructed from type variables, typical representative α, primitive (ground) types, typical representative ι, and the function space constructor:

$$\tau ::= \alpha \mid \iota \mid \tau \to \tau$$

The algorithm will produce the principal type scheme for a term; type schemes have the following form:

$$\sigma ::= \tau \mid \forall\alpha.\sigma$$

We will use the shorthand $\forall\alpha_1 \ldots \alpha_n.\sigma$ for $\forall\alpha_1 \ldots \forall\alpha_n.\sigma$; the α_i are called *generic type variables*. A *monotype* is a type containing no type variables.

A substitution is a mapping from type variables to types. For a substitution S, we write:

$$S\sigma$$

to represent the type scheme obtained from σ by replacing each free occurrence of any variable in the domain of S by the corresponding element of the co-domain of S; the resultant type scheme is called an *instance* of σ. We sometimes write S explicitly as:

$$[\tau_1/\alpha_1, \ldots, \tau_n/\alpha_n]$$

meaning that τ_i $(1 \leq i \leq n)$ is substituted for α_i. Notice that the substitution operation may lead to variable capture if applied naively — we should adopt a variable convention.

In contrast to the notion of instance, a type scheme $\sigma = \forall\alpha_1 \ldots \alpha_m.\tau$ has a *generic instance* $\sigma' = \forall\beta_1 \ldots \beta_n.\tau'$ if $\tau' = [\tau_i/\alpha_i]\tau$ and the β_j are not free in σ; in this case we write $\sigma > \sigma'$. Notice that instantiation involves substitution for free variables while generic instantiation acts on bound variables.

We now present a formal system for type inference. The basic judgements, or assertions, in this system are of the form:

$$A \vdash e : \sigma$$

where A is a set of assumptions of the form:

$$x : \sigma' \text{ where } x \text{ is a variable}$$

The assertion should be read: "Under assumptions A, e has type σ". An assertion is closed if A and σ contain no free variables. The axioms and the rules are presented in Figure 4.

The assumptions A_x used in **Abs** and **Let** denote the new assumptions derived from A by removing any assumption about x. The reader should compare these rules, particularly **Comb** and **Abs**, to the definition of λ^τ-terms. Notice that polymorphism is represented by type schemes; only the rules **Taut**, **Inst**, **Gen** and **Let** concern type schemes. Type inference amounts to a process of theorem proving in this formal system, for example:

$$\dfrac{\dfrac{x : \alpha \vdash x : \alpha \quad \textbf{Taut}}{\vdash (\lambda x.x) : \alpha \to \alpha} \ \textbf{Abs}}{\vdash (\lambda x.x) : \forall\alpha.\alpha \to \alpha} \ \textbf{Gen}$$

This (polymorphic) type associated with the identity function is the *most general type* for the identity function; all other possible types are generic instances of $\forall\alpha.\alpha \to \alpha$ – it is the largest type in the $>$-ordering.

We now present an algorithm for inferring types; the algorithm is Milner's $\mathcal{W}$ algorithm. The informal type of $\mathcal{W}$ is:

$$\textbf{Taut} \quad A \vdash x : \sigma \qquad\qquad (x : \sigma \text{ in } A)$$

$$\textbf{Inst} \quad \frac{A \vdash e : \sigma}{A \vdash e : \sigma'} \qquad\qquad (\sigma > \sigma')$$

$$\textbf{Gen} \quad \frac{A \vdash e : \sigma}{A \vdash e : \forall \alpha . \sigma} \qquad\qquad (\alpha \text{ not free in } A)$$

$$\textbf{Comb} \quad \frac{A \vdash e : \tau' \to \tau \qquad A \vdash e' : \tau'}{A \vdash (e\ e') : \tau}$$

$$\textbf{Abs} \quad \frac{A_x \cup \{x : \tau'\} \vdash e : \tau}{A \vdash (\lambda x . e) : \tau' \to \tau}$$

$$\textbf{Let} \quad \frac{A \vdash e : \sigma \qquad A_x \cup \{x : \sigma\} \vdash e' : \tau}{A \vdash (let\ x = e\ in\ e') : \tau}$$

Figure 4. Polymorphic type inference.

$$\text{Assumptions} \times \text{Exp} \to \text{Substitution} \times \text{Type}$$

and if:

$$\mathcal{W}(A, e) = (S, \tau)$$

then we have:

$$SA \vdash e : \tau$$

where substitutions are extended to assumption lists in the obvious way. In order to define $\mathcal{W}$, we will need two operations: unification and closure with respect to some assumptions.

DEFINITION 99. A *unifier* of two terms is a substitution which, when applied to the two terms, makes the terms equal. We will define an algorithm $\mathcal{U}$ which finds a unifier for two types τ and τ' or fails. Furthermore:

1. If $\mathcal{U}(\tau, \tau') = V$ then $V\tau = V\tau'$
   i.e. V unifies τ and τ'
2. If S unifies τ and τ' then $\mathcal{U}(\tau, \tau')$ returns some V and there is another substitution R such that

   $$S = RV$$

   where composition of substitutions is done in the obvious way. This requirement amounts to stating that V does the least amount of work to equate the two terms; V is called the *most general unifier*.

3. V only involves variables in τ and τ'; no new variables are introduced during unification.

The algorithm uses the notion of a *disagreement set*:

$$\begin{array}{lll} D(\tau, \tau') & = & \emptyset \\ & & \text{if } \tau = \tau' \\ & = & \{(\tau_1, \tau_1')\} \\ & & \text{if } \tau_1, \tau_1' \text{ are the ``first'' two subterms at which } \tau \text{ and } \tau' \text{ disagree} \end{array}$$

In the second clause of the definition, we assume a depth-first traversal. Some examples may clarify this concept:

$$\begin{array}{rcl} D(int \to int, int \to int) & = & \emptyset \\ D(\alpha \to \beta, \alpha \to \beta) & = & \emptyset \\ D(\alpha, \alpha \to \beta) & = & \{(\alpha, \alpha \to \beta)\} \\ D(\alpha \to \alpha, (int \to int) \to \beta) & = & \{(\alpha, int \to int)\} \\ D((int \to \alpha) \to \beta, (int \to int) \to \gamma) & = & \{(\alpha, int)\} \end{array}$$

We now define $\mathcal{U}$ in terms of an auxiliary function which iterates with a substitution and the two types to find the unifier:

$$\begin{array}{lll} \mathcal{U}(\tau, \tau') & = & iterate(Id, \tau, \tau') \\ \text{where} & & \\ iterate(V, \tau, \tau') & = & \text{if } V\tau = V\tau' \\ & & \text{then } V \\ & & \text{elsif } a \text{ is a variable that does not occur in } b \\ & & \text{then } iterate([b/a]V, \tau, \tau') \\ & & \text{elsif } b \text{ is a variable that does not occur in } a \\ & & \text{then } iterate([a/b]V, \tau, \tau') \\ & & \text{else FAIL} \\ & & \text{where } \{(a, b)\} = D(V\tau, V\tau') \end{array}$$

The closure of a type results in a type where some free variables are quantified; more formally:

DEFINITION 100. The *closure* of a type τ with respect to some assumptions A involves making any free variables of τ which are not free in A into generic type variables. We write the closure as $\overline{A}(\tau)$. Thus:

$$\overline{A}(\tau) = \forall \alpha_1 \dots \alpha_n . \tau$$

where $\alpha_1, \dots, \alpha_n$ are the type variables occurring free in τ but not in A.

We define $\mathcal{W}$ in Figure 5.

$\mathcal{W}(A, e) = (S, \tau)$ where

- If $e \equiv x$ and $x : \forall\alpha_1 \ldots \alpha_n.\tau' \in A$ then $S = Id$ and $\tau = [\beta_i/\alpha_i]\tau'$ with the β_i new.

- If $e \equiv e_1 e_2$:
  let $\mathcal{W}(A, e_1) = (S_1, \tau_1)$ and
  $\mathcal{W}(S_1 A, e_2) = (S_2, \tau_2)$
  and $\mathcal{U}(S_2\tau_1, \tau_2 \to \beta) = V$ where β is new
  then $S = VS_2S_1$ and $\tau = V\beta$.

- If $e \equiv \lambda x.e_1$:
  let β be a new type variable and $\mathcal{W}(A_x \cup \{x : \beta\}, e_1) = (S_1, \tau_1)$
  then $S = S_1$ and $\tau = S_1\beta \to \tau_1$.

- If $e \equiv let\ x = e_1\ in\ e_2$:
  let $\mathcal{W}(A, e_1) = (S_1, \tau_1)$ and
  $\mathcal{W}(S_1 A_x \cup \{x : \overline{S_1 A}(\tau_1)\}, e_2) = (S_2, \tau_2)$
  then $S = S_2 S_1$ and $\tau = \tau_2$.

- Otherwise $\mathcal{W}$ fails.

Figure 5. The algorithm $\mathcal{W}$.

We now state several important properties of this algorithm.

PROPOSITION 101. *If S is a substitution and $A \vdash e : \sigma$ holds then $SA \vdash e : S\sigma$ also holds. Moreover if there is a derivation of $A \vdash e : \sigma$ of height n then there is also a derivation of $SA \vdash e : S\sigma$ of height less than or equal to n.*

THEOREM 102 (Soundness of $\mathcal{W}$).
If

$$\mathcal{W}(A, e) = (S, \tau)$$

then

$$SA \vdash e : \tau$$

which is just the property that we required in the specification of $\mathcal{W}$.

Given A and e, σ_p is a *principal type scheme* of e under A if and only if:

- $A \vdash e : \sigma_p$
- Any other σ for which $A \vdash e : \sigma$ is a generic instance of σ_p.

Soundness states (approximately) that any types inferred by $\mathcal{W}$ can be inferred using the inference system. An equally important property is *completeness*, which means that any type that can be inferred by the inference system can be found by $\mathcal{W}$ (again approximately). We state two versions of completeness without proof; see the seminal paper by Damas and Milner for details.

1. **Completeness of $\mathcal{W}$:**

   Given A and e, let A' be an instance of A and σ a type scheme such that $A' \vdash e : \sigma$ then:

   - $\mathcal{W}(A, e)$ succeeds
   - If $\mathcal{W}(A, e) = (S, \tau)$ then for some substitution R:

   $$A' = RSA$$

   and $R\overline{SA}(\tau) > \sigma$

2. **Completeness (no free type variables) of $\mathcal{W}$:**

   If $A \vdash e : \sigma$, for some σ, then $\mathcal{W}$ computes a principal type scheme for e under A.

Property 2 is actually a simple corollary of Property 1.

We conclude this section by noting the importance of the let-construct in the language. In a type-free setting:

$$let\ x = e\ in\ e'$$

is just syntactic sugar for:

$$(\lambda x.e')e$$

This is no longer true when we use $\mathcal{W}$ to infer types. The let-construct introduces polymorphic functions; thus:

$$let\ f = \lambda x.x\ in \ldots f\ true \ldots f\ 1 \ldots$$

can be validly typed because f will be given type $\forall\alpha.\alpha \rightarrow \alpha$ (because of the closure operator in the fourth clause) which can then be instantiated to both $bool \rightarrow bool$ and $int \rightarrow int$. However:

$$(\lambda f. \ldots f\ true \ldots f\ 1 \ldots)(\lambda x.x)$$

cannot be validly typed because $\lambda x.x$ is given type $\alpha \rightarrow \alpha$ and α can only be instantiated to one type.

7.4 Intersection Types

In the polymorphic λ-calculus a function can be applied to arguments of different types but the types of the arguments must have the same "structure". This becomes more apparent in the context of programming languages where we have a richer set of type constructors. For example, the standard *map* function is polymorphic in its second argument but the argument must at least be a list structure. Most programming languages also allow overloading of operators, for example + can be applied to a pair of integers or a pair of reals — the operation performed in each case is very different. In terms involving overloaded operators, functions are applied to arguments with structurally different types.

An *intersection* type is like a type in the simply typed calculus except types can be constructed using the intersection operator $\cap$ — a term which is assigned such a type has both types involved in the intersection. An example of how this is used is in the term $\lambda x.xx$, which is untypable in the previous calculi, but we can show:

$$(\lambda x.xx) : (\sigma \cap (\sigma \rightarrow \tau)) \rightarrow \tau$$

Notice that the argument is given both σ and $\sigma \rightarrow \tau$ as types and thus the self-application in the body can be typed. In this section, we will present the $\lambda\cap$-calculus. In this calculus, it no longer makes sense to have explicit typing so we present an implicitly typed calculus.

The set of types is defined as follows:

$$\tau ::= \alpha \mid \iota \mid \tau \rightarrow \tau \mid \tau \cap \tau$$

Amongst the constants, we include a distinguished type ω.

$$\sigma \leq \sigma$$

$$\sigma \leq \omega \qquad \omega \leq \omega \to \omega$$

$$\sigma \cap \tau \leq \sigma \qquad \sigma \cap \tau \leq \tau$$

$$\frac{\sigma \leq \tau \quad \tau \leq \rho}{\sigma \leq \rho} \qquad \frac{\sigma \leq \tau \quad \sigma \leq \rho}{\sigma \leq \tau \cap \rho}$$

$$(\sigma \to \rho) \cap (\sigma \to \tau) \leq \sigma \to (\rho \cap \tau)$$

$$\frac{\sigma' \leq \sigma \quad \tau \leq \tau'}{(\sigma \to \tau) \leq (\sigma' \to \tau')}$$

Figure 6. The pre-order on intersection types.

The rôle of ω is as a *universal* type; any term can be assigned ω as a type. Given ω and $\cap$ it is fairly natural to order the types; we define the pre-order in Figure 6.

We write $\sigma \equiv \tau$ in the case that $\sigma \leq \tau$ and $\tau \leq \sigma$. We also adopt the convention that $\cap$ has higher precedence than $\to$ which allows us to omit some parentheses.

The last rule in the definition of $\leq$ expresses the fact that $\to$ is *contravariant* in its first argument.

The inference system shown in Figure 7 assigns intersection types to terms.

In $\lambda\cap$, the following properties hold:

- $\beta\eta$ is CR.
- SN fails – every term from Λ is typable, including Ω; all terms have type ω.
- it is undecidable whether a term has a particular type.

van Bakel has studied a restricted inference system which does not have the **Top** rule; in this system, the following is true:

$$SN(M) \Leftrightarrow \exists A.\exists \sigma.A \vdash M : \sigma$$

Barendregt *et al* prove that:

$$M \text{ has a normal form} \Leftrightarrow \exists A.\exists \sigma.A \vdash M : \sigma \text{ and } \omega \text{ does not occur in } \sigma$$

$$\textbf{Taut} \quad A \vdash x : \sigma \qquad\qquad (x : \sigma) \in A$$

$$\textbf{Top} \quad A \vdash M : \omega$$

$$\rightarrow\!\textbf{E} \quad \frac{A \vdash M : (\sigma \rightarrow \tau) \quad A \vdash N : \sigma}{A \vdash M\ N : \tau} \qquad \rightarrow\!\textbf{I} \quad \frac{A_x \cup (x : \sigma) \vdash M : \tau}{A \vdash \lambda x.M : \sigma \rightarrow \tau}$$

$$\cap\textbf{E} \quad \frac{A \vdash M : (\sigma_1 \cap \sigma_2)}{A \vdash M : \sigma_i}\ i = 1, 2 \qquad \cap\textbf{I} \quad \frac{A \vdash M : \sigma \quad A \vdash M : \tau}{A \vdash M : \sigma \cap \tau}$$

$$\textbf{Sub} \quad \frac{A \vdash M : \sigma \quad \sigma \leq \tau}{A \vdash M : \tau}$$

Figure 7. Type inference for intersection types.

8 CONCLUSIONS

We have given a Computer Science perspective on the material covered in this article. Many of the fundamental results have been produced by logicians. Our inspiration has been Barendregt's encyclopaedic tome [Barendregt, 1984]. Most of the basic material presented here is treated in much more detail in [Barendregt, 1984] and the interested reader is urged to consult it. The classical presentation of the λ-calculi is Church's 1941 report [Church, 1941].

One of the earliest textbooks in the field was Hindley, Lercher and Seldin's *Introduction to Combinatory Logic*; while this is no longer available, [Hindley and Seldin, 1986] is a much-expanded treatment of the same material. The latter is written from a mathematical logician's perspective, so it is short on computational intuitions, but it is nonetheless a useful reference, particularly for material on typed calculi. Krivine's book, [Krivine, 1993], covers similar material but includes a chapter on intersection types. The type free λ-calculus is the main object of study in [Hankin, 1994]; this book is written from a computer science perspective and explores some of the links between the λ-calculus and programming language theory.

A number of functional programming textbooks contain computing-oriented descriptions of the λ-calculus and combinators. For example [Field and Harrison, 1988; Peyton Jones, 1987; Reade, 1989] contain accounts of the main results relating to functional languages and their implementation. The de Bruijn notation is studied in detail in [Curien, 1993]. Michaelson uses the λ-calculus to introduce functional programming in [Michaelson, 1989]. Turner, [Turner, 1991] provides an introduction to the axiomatic foundations of functional programming which includes a lot of material on both the λ-calculus and type theory.

A more detailed consideration of evaluation strategies for functional languages and the relevance of the λ-calculus to such languages may be found in [Field and Harrison, 1988]. The material on labelled reduction and residuals is based on the approach used by Klop in [Klop, 1980].

A comprehensive treatment of semantics may be found in [Barendregt, 1984] and [Hindley and Seldin, 1986]. Both include detailed discussions of Scott's models. Stoy's book [Stoy, 1977] is the classical textbook on denotational semantics and contains a good introduction to the λ-calculus and models. There have been a number of books published which contain some coverage of this material; a good example is [Winskel, 1993].

Computability aspects of the Lambda Calculus are dealt with in [Barendregt, 1984] and [Hindley and Seldin, 1986]. A more general treatment of this subject may be found, for example, in [Hopcroft and Ullman, 1979].

In [Barendregt, 1984] the main focus is on type-free calculi; there is a short appendix on the simple typed λ-calculus. For a more detailed treatment of typed calculi, [Barendregt, 1992] is recommended. A large part of [Hindley and Seldin, 1986] is devoted to typed calculi and [Huet, 1990] contains a number of seminal papers on the polymorphic λ-calculus. The material on Milner's algorithm is based on [Damas and Milner, 1982].

As mentioned in the Introduction, an interesting use of λ-calculus has been in natural language understanding; the book [Dowty *et al.*, 1981] is a good introduction to Montague semantics.

In constructing this "bibliography" we have restricted our attention to material that is readily accessible; except in a few cases, this has meant that we have cited books. Many of the most fundamental and exciting results have appeared and continue to appear in conference proceedings and journals. Good starting points are the proceedings of the ACM Symposium on *Principles of Programming Languages* (POPL), the IEEE Symposium on *Logic in Computer Science* (LICS) and the proceedings of the ETAPS Federated conference.

BIBLIOGRAPHY

[Barendregt, 1984] H. P. Barendregt, *The Lambda Calculus: Its Syntax and Semantics*, 2nd edition, North Holland, 1984.

[Barendregt, 1992] H. P. Barendregt, Lambda Calculi with Types, in *Handbook of Logic in Computer Science, Volume II*, S. Abramsky, D. Gabbay and T. S. E. Maibaum (eds), Oxford University Press, 1992.

[Church, 1941] A. Church, *The Calculi of Lambda Conversion*, Princeton University Press, 1941.

[Curien, 1993] P.-L. Curien, *Categorical Combinators, Sequential Algorithms and Functional Programming*, 2nd edition, Birkhäuser, 1993.

[Damas and Milner, 1982] L. Damas and R. Milner, Principal Type Schemes for Functional Programs, in *Proceedings of POPL'82*, ACM Press, 1982.

[Dowty *et al.*, 1981] D. R. Dowty, R. E. Wall and S. Peters, *Introduction to Montague Semantics*, D. Reidel Publishing Company, 1981.

[Field and Harrison, 1988] A. J. Field and P. G. Harrison, *Functional Programming*, Addison Wesley, 1988.

[Hankin, 1994] C. L. Hankin, *Lambda Calculi: A Guide for Computer Scientists*, Oxford University Press, 1994.

[Hindley and Seldin, 1986] J. R. Hindley and J. P. Seldin, *Introduction to Combinators and λ-Calculus*, Cambridge University Press, 1986.

[Hopcroft and Ullman, 1979] J. E. Hopcroft and J. D. Ullman, *Introduction to Automata Theory, Languages and Computation*, Addison Wesley, 1979.

[Huet, 1990] G. Huet, *Logical Foundations of Functional Programming*, Addison Wesley, 1990.

[Klop, 1980] J. W. Klop, *Combinatory Reduction Systems*, CWI Report, Amsterdam, 1980.

[Krivine, 1993] J. L. Krivine, *Lambda-Calculus, Types and Models*, Masson, 1993.

[Michaelson, 1989] G. Michaelson, *An Introduction to Functional Programming Through Lambda Calculus*, Addison Wesley, 1989.

[Peyton Jones, 1987] S. L. Peyton Jones, *The Implementation of Functional Programming Languages*, Prentice Hall International, 1987.

[Reade, 1989] C. Reade, *Elements of Functional Programming*, Addison Wesley, 1989.

[Stoy, 1977] J. E. Stoy, *Denotational semantics: the Scott–Strachey approach to programming language theory*, MIT Press, 1977.

[Turner, 1991] R. Turner, *Constructive Foundations for Functional Languages*, McGraw Hill, 1991.

[Winskel, 1993] G. Winskel, *The Formal Semantics of Programming Languages*, MIT Press, 1993.

DOV M. GABBAY AND LARISA L. MAKSIMOVA

INTERPOLATION AND DEFINABILITY

This chapter is on interpolation and definability. This notion is not only central in pure logic, but has significant meaning and applicability in all areas where logic itself is applied, especially in computer science, artificial intelligence, logic programming, philosophy of science and natural language. The notion may sometimes appear to the reader as too technical/mathematical but it does also have a general meaning in terms of expressibility and definability.

Many of basic results and methods on the subject are presented in our book [Gabbay and Maksimova, 2005] published in Oxford University Press. The aim of this chapter is to explain the various options and aspects of interpolation and to give some case study examples (for the benefit of the applied reader who does not wish to read the book in its entirety) and then to give an overview of basic results on interpolation and definability in non-classical, especially in modal and intuitionistic logics.

1 INTRODUCTION AND DISCUSSION

1.1 General discussion

Let us begin with the simplest notion of interpolation. Let $\vdash$ be a consequence relation for a (propositional) logic based on the atomic propositions $\{q_1, q_2, \ldots\}$. The reader is invited to think of classical or intuitionistic propositional logics as examples. Let $A(p, q)$ and $B(q, r)$ be two formulas built up from the connectives of the logic and the atoms $\{p, q\}$ and $\{q, r\}$ respectively. So A and B have (the language based on) the atom $\{q\}$ in common.

Assume now that we have $A \vdash B$. The simple interpolation property asserts that there exists a formula $H(q)$ built up from the atoms of the common language $\{q\}$ such that $A \vdash H$ and $H \vdash B$ hold.

At first sight this assertion seems simple and straightforward. However, let us examine it more closely, by offering a variety of points of view. Note that we mentioned interpolation for atomic propositions but one can also discuss interpolation for connectives. For example let $\Box_1, \Box_2, \Box_3$ be three modalities (for example, necessity, knowledge and time operators). Assume $A(\Box_1, \Box_2) \vdash B(\Box_2, \Box_3)$ is there an $H(\Box_2)$ built up using $\Box_2$ only such that $A \vdash H$ and $H \vdash B$?

The following is a specific problem: Let $K \otimes K \otimes K$ be some product of propositional modal K with itself (there are various possibilities for $\otimes$ see [Gabbay *et al.*, 2003]). Assume $A(p, q, \Box_1, \Box_2) \vdash B(q, r, \Box_2, \Box_3)$. Is there an interpolant $H(q, \Box_2)$ such that $A \vdash H \vdash B$?

D. Gabbay and F. Guenthner (eds.),
Handbook of Philosophical Logic, Volume 15, 67–123.
DOI 10.1007/978-94-007-0485-5_2, © Springer Science+Business Media B.V. 2011

Interpolation: View 1 — Common logical content

This view is most plausible in applications. If A "talks" or "specifies" properties of $\{p, q\}$ and these properties "force" properties involving $\{q, r\}$, then since q is the only common factor, we would expect that there must be something about $\{q\}$, namely $H(q)$, which is forced by A, i.e., $A \vdash H$ and this "requirement" $H(q)$ is the one which forces B, i.e., $H \vdash B$.

To give an example, suppose the Queen of England invites several famous professors of logic for dinner in her palace. She specifies that the *table* (q) and *chairs* $(p_1, p_2, \ldots)$ are arranged in such a way that the result is completely symmetrical, (call this specification), $S(q, p_i)$. She also specifies that the *tablecloth* (r) must lie on the table in a completely flawless and smooth way (call this specification) $F(q, r)$.

The royal staff discover that $S\&F \vdash$ *tablecloth is round.*

Interpolation in this case means that for some $H(q)$

$$S \vdash H(q) \vdash F \rightarrow \textit{ tablecloth is round}.$$

H says in this case that the table itself must be round and this forces the tablecloth to be round.

Interpolation: View 2 — Expressive power

We immediately encounter some technical difficulties with our simple view of interpolation. The first is simple. Take the fragment of logic with implication $\rightarrow$ only. Then we have

$$p \rightarrow p \vdash r \rightarrow r$$

The interpolant is $\top$ but technically it is not in the language. Similarly we can have

$$p\&\neg p \vdash q\&\neg q$$

and then we would need $\perp$ in the language.

The reader might say that these are artificial examples of lack of interpolation and that $\top$ and $\perp$ should be in the language anyway. However, this is not a satisfactory answer. First, there are logics where the connectives are not necessarily expected to be available. Second, there may be another way of looking at interpolation for such logics, namely as forcing *functional completeness.*

Consider, for example Łukasiewicz three-valued logic for $\rightarrow, \&, \vee, \neg$ with 0 as the designated value and the table of Figure 1.

The value $\frac{1}{2}$ cannot be defined by a formula (i.e., a formula $A(q_1, \ldots, q_n)$ which gets value $\frac{1}{2}$ for all $(q_1, \ldots, q_n)$).

Define $A(p, q) \vdash B(q, r)$ iff for all values of p, q, r we have

$$\text{value } A \geqslant \text{value } B.$$

We have

$$p\&\neg p \vdash q \vee \neg q$$

but no interpolant (we need $\frac{1}{2}$ as interpolant). It is much less obvious in this case that a constant for $\frac{1}{2}$ is needed in the language. We might as well argue that any finite many valued matrix logic must have constants for all its values.

p	q	$p \vee q$	$p\&q$	$p \to q$	$\neg p$
0	0	0	0	0	1
0	$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{2}$	1
0	1	0	1	1	1
$\frac{1}{2}$	0	0	$\frac{1}{2}$	0	$\frac{1}{2}$
$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{2}$
$\frac{1}{2}$	1	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$
1	0	0	1	0	0
1	$\frac{1}{2}$	$\frac{1}{2}$	1	0	0
1	1	$\frac{1}{2}$	1	0	0

Figure 1.

However, we notice that requiring interpolation may be equivalent to functional completeness in many logics because it forces us to have constants for truth values, and the functions associated with the other connectives of the logic together with the constant names of the values may give us functional completeness.

Note that in some cases, such as $p \to p \vdash r \to r$, we can try and formulate interpolation slightly differently. If $A, B \vdash r_A$ (r_A atomic in the language of A), then for some H in the common language, (maybe we don't need H at all), we have

(1) $B \vdash H$,

(2) $A, H \vdash r_A$.

In the case above we write $r, p \to p \vdash r$ and thus drop $p \to p$.

In the general case interpolation is genuinely linked to expressive power and the lack of connectives cannot be dismissed so easily. A more technical example of lack of interpolation, because a connective is missing, can be found in [Gabbay and Maksimova, 2005, Chapter 15], where the existential quantifier "$\exists x$" is needed to interpolate, but is missing.

In fact, the lack of interpolation because of the lack of expressibility can be quite sophisticated mathematically. The reader can have his own opinion of whether technical examples like the above are genuine cases of lack of interpolation. The mathematician will probably be quite happy with an exhaustive classification of systems with and without interpolation and will take a special pride and delight in the various methods and counterexamples used. The computer scientist, however, may take a different view. He may not want a "laundry-list" of variations on meaningless technical results arising from an "obvious" lack of expressive power. He may want to know the reasons behind the results. The truth is that the mathematician would also like to understand what is happening. It seems that there are serious cases where lack of expressive power has meaning beyond the technical. See, for example, [Fine, 1979], where it is shown that interpolation fails for quantified modal S5. Section 14.3 of [Gabbay and Maksimova, 2005] discusses the expressibility reasons for the failure. A long history and hard work of many colleagues are involved here over a period of 20 years.

Let us look at another example, from linear temporal logic. There is lack of interpolation there because a connective "until" is definable using additional propositional atoms but it is not definable without them. In this example technical expressive power has an intuitive meaning. Temporal logic is applied in the analysis of tenses of natural language, and from the linguistic point of view, expressing temporal phenomena using additional atoms (atoms correspond to time dates like $p =$ "The date is January 1st, 1970") is *not* the same kind of linguistic construction as expressing temporal phenomena using connectives (temporal connectives correspond to tense and aspect constructions in natural language). So lack of interpolation for the linguist may carry a whole lot more meaning than lack of interpolation may mean for the computer scientist!

We will come back to tenses and time later on in this section, but meanwhile, let us continue to our next view.

Interpolation: View 3 — Quantifier elimination

Consider the situation of $A(p, q) \vdash H(q)$.

Interpolation has nothing to say about this situation, since H is in a sublanguage of A. However, the computer scientist has many questions to ask. If A is a specification about $\{p, q\}$ and it forces a property $H(q)$ of q, we ask which part of A does the job? In other words, which part of A is "responsible for" q?

This is important to the computer scientist for several reasons.

1. Can the specification A be algorithmically simplified to give the part dealing with q only? (Imagine $A(p, q)$ is to be implemented using gates and boolean circuits, we want to simplify the circuitry to something controlling q only.)

2. Suppose A is a program that does not perform exactly right. Can we *slice* it (the area doing this kind of work is called "*program slicing*") to the parts on $\{p\}$ and $\{q\}$ and check where the bug is?

Another example is when a specification $H(q_i)$ is implemented in another computer language and the implementation is described by the formula $A(q_i, p_j)$. (Say, implementing sets in the programming langauge C as lists). If the implementation is *sound*, we want $A(q_i, p_j) \vdash H(q_i)$. Again we want to know exactly how A is "responsible" for H and most importantly, what additional properties and expressive power does A *add* to H.

Put in technical terms, we want to effectively derive from $A(p, q)$ the part $A \restriction \{q\}$, talking about q only. In logical terms, we want to have

$$A \vdash A \restriction \{q\}$$

and we want $A \restriction \{q\}$ to be minimal, i.e., whenever $A \vdash H(q)$ then $A \restriction \{q\} \vdash H(q)$.

Put in these words, this has meaning in pure logic as well. Obviously logically $A \restriction \{q\}$ is equivalent to $\exists p A(p, q)$ and the computational requirements of the "slicing" computer scientist amount to (propositional and/or second order) quantifier elimination in logic.

In fact, if

$$A(p, q) \vdash B(q, r)$$

and the deduction theorem is available and the logic $\vdash$ is reasonable, then we can write schematically:

$$\vdash \forall p \forall r (A(p, q) \to B(q, r)),$$
$$\exists p A(p, q) \vdash \forall r B(q, r)$$

and hence

$$A(p, q) \vdash \exists p A(p, q) \vdash \forall r B(q, r) \vdash B(q, r).$$

In classical logic we can write

$$\exists p A(p, q) \equiv A(\top, q) \vee A(\bot, q)$$

which gives us an algorithmic way of finding interpolants. This can be done in any finite fixed matrix logic (for example, finite many valued logic) provided constants for the values and "disjunction" are available.

Interpolation: View 4 — Artificial intelligence

The quantifier elimination view is important for interpolation for non-monotonic logics arising in Artificial Intelligence. Most of the logics there, such as default logics, circumscription, logics with negation as failure, inheritance nets, autoepistemic logics and more are nonmonotonic and nontransitive. We consider these logics as very applicable common sense, day to day service logics which *allow* for the following to happen:

(1) $\Delta \vdash B$ but $\Delta, A \nvdash B$

(2) $A \vdash B$ and $B \vdash C$ but $A \nvdash C$

Clearly we have a difficulty here for interpolation. If $A(p, q) \vdash B(q, r)$, we still expect some interpolant $H(q)$ to "transmit" the specification from A to B via $A \vdash H \vdash B$. But if such logics cannot force transitivity, how can we expect this?

The existential quantifier view is more helpful. $\exists p A(p, q)$ has nothing to do with B and is likely to have a similar meaning in nonmonotonic logic, parallel to the monotonic case.

Interpolation for nonmonotonic logics in AI have not been studied much. Partly because the logics have neither good proof theory nor simple semantics and partly because we need to reconsider what interpolation is supposed to mean in such cases.

Note that it is not clear what the existential view would be for connective interpolation of the form

$$A(\Box_1, \Box_2) \vdash B(\Box_2, \Box_3).$$

Interpolation: View 5 — Proof theory

The "slicing" view of interpolation has its counterpart in logic. It is the proof theoretical view for interpolation. Suppose we are given a logic $\vdash$ and we want to find the interpolants algorithmically in this logic, then we must ask, as logicians, how is $\vdash$ presented to us? Let us assume that we have some proof theoretic manipulative rules that can begin with $A(p, q)$ and manipulate it and end up with $B(q, r)$. Thus $A(p, q) \vdash B(q, r)$ is established algorithmically. We can thus hope, by carefully and inductively looking at the proof and the rules, to extract an $H(q)$ such that $A(p, q) \vdash H(q) \vdash B(q, r)$.

In fact a careful analysis of the proof process may give us additional information about the interpolant H. We can indicate the positions of q in H as compared with its position in A and the nature and nesting of the connectives in H as compared with A. In practice it is not so difficult to find the interpolant once a good set of proof rules is established. The hard part to prove is that a set of good rules does indeed characterise the logic. The reader has probably heard a lot about cut free proofs and normalisation theorems, as well as goal directed procedures.

The basic idea is that if we start with some assumptions $A_1, \ldots, A_n$, we first break them apart into smaller subparts (using elimination rules) and then put the parts together again (using introduction rules). Such a proof is *normalised* and it is easy to trace when and where the new (non-logical) symbols are introduced.

This view seems very promising indeed. In fact it has three further major advantages.

1. Proof methodologies go across logics: variations in the proof process can yield a variety of known logics. Thus interpolation can be studied for families of logics and hence its nature can be better understood (see [Gabbay and Olivetti, 2000]).

2. Logicians love cut free formulations of logics and so one can obtain a lot of support for interpolation in many quarters. In fact we hope that this chapter will contribute towards crystallising some interest in our area.

3. In computer science the theme "proofs as programs" is a very prominent, strongly supported and widely applied approach. Computer science is also keenly interested in interpolation. So if we study some logics (usually intuitionistic or linear logic variants) which also play a role also as programming languages, we can unify and benefit from both points of view in our quest of understanding interpolation. To succeed in this attempt we need to look at some simple logic programming language (intuitionistic logic fragment with negation as failure? Horn clause logic?) and try and prove interpolation for it in a variety of ways.

Interpolation: View 6 — Consistency

The consistency view is important to both the computer scientist and the logician. Suppose Δ and Γ are two specifications/theories which "agree" on their common language. We ask ourselves can we put them together in a consistent way, i.e., consider $\Theta = \Delta \cup \Gamma$, and is Θ consistent? This is known in logic as Robinson's consistency theorem (for different logics there may be different specific formulations and assumptions involved in the proof of this theorem). In computer science this is the problem of amalgamation (or push-out to use category theory language).

Let $H = \Delta \cap \Gamma$. Then Figure 2 describes the situation.

We seek a commutative diagram here. Figure 2 is meaningful in computer science. We can give different groups of programmers the tasks to write specification and codes on different parts of the application. We hope they agree on the common parts and we hope we can put them all together consistently.

The connection with interpolation is clear. If $\Delta \cup \Gamma$ were inconsistent then $\Delta, \Gamma \vdash \perp$. We find an interpolant I in the common language such that $\Gamma \vdash I$ and $\Delta, I \vdash \perp$, i.e., $\Gamma \vdash I$ and $\Delta \vdash I \to \perp$. But this contradicts the assumption that Δ and Γ agree on the common language. The above "proof" manipulated Δ, Γ and $\perp$ in a way permissible in classical logic. When we deal with a variety of logics we get several possible formulations of the interpolation and of the consistency theorems and one has to study

$$
\begin{tikzcd}
& \Theta & \\
\Delta \arrow[ur] & & \Gamma \arrow[ul] \\
& H \arrow[ul] \arrow[ur] &
\end{tikzcd}
$$

Figure 2.

what implies what. See for example Sections 2.1 and 2.3 below.

In fact, the very notion of "consistency" can vary from logic to logic (as well as the availability of $\perp$) and it is better to talk about "Θ is acceptable" instead of "Θ is consistent." The notion of "acceptability" is certainly more in tune with artificial intelligence applications, where theories may be formally logically consistent but unacceptable in the application area. For example, any database which does not satisfy its *integrity constraints* is consistent but not acceptable.

Interpolation: View 7 — Semantical view

The consistency view leads naturally to semantical and algebraical considerations. Roughly speaking, given $A \vdash B$ we can proceed as follows: Let $\mathbb{H}$ be $\mathbb{H} = \{h \text{ in common language } \mid A \vdash h\}$. If $\mathbb{H} \vdash B$ then we have our interpolant. Otherwise $\Delta = \{A\} \cup \mathbb{H}$ and some carefully chosen $\Gamma \supseteq \{B \to \perp\}$ are each consistent and they agree on the common language. Use a Henkin or tableaux construction to build a model for $\Delta \cup \Gamma$ (consistency theorem) and get a contradiction.

This method is powerful but it has its problems.

1. Many logics have models involving families of interconnected theories (possible worlds) and so it is complicated to construct a common model for $\Delta \cup \Gamma$ out of Δ and Γ.

2. We already saw that sometimes interpolation fails because of lack of expressive power. If the attempted construction of a model for $\Delta \cup \Gamma$ fails, can we get an idea of what expressiveness is missing? It is not as "easy" in the semantic view as it might be in the proof theoretic view.

Some logics such as intuitionistic logic of constant domains (for which the problem of interpolation is still open) has been giving us a lot of headache over the past 30 years, and several false proofs using different methods have been published.

Note that the logic or the theory for which we want to interpolate may be presented semantically to us via a class of models $\mathcal{M}$. Thus interpolation will take the form $\mathcal{M} \vDash A \to B$ implies that for some H in the common language $\mathcal{M} \vDash A \to H$ and $\mathcal{M} \vDash H \to B$. In such a case only semantic methods are available.

Interpolation: View 8 — Algebraic view

This is one of the most potent views for interpolation. Given a logic $\vdash$ it must be such that a Lindenbaum free algebraic semantics is available for it. The atoms are the generators of the algebra, the connectives are the algebraic functions and the axioms are identities. Any further axioms are additional identities. We can ask whether such algebras (for a given logic) satisfy a variety of amalgamation properties. Interpolation is connected with amalgamation and because of the effectiveness and potency of algebraic methods, impressive results can be obtained. See Section 2.3 for example.

Note that amalgamation is a general algebraic property of independent research interest.

The amalgamation view is connected with the consistency and semantic views. A close examination of the constructions will show that one is actually building models. There is also a connection with amalgamation of proof systems (called combining logics).

It is worth our while to look more closely at combining logics. Let $\mathbf{L}_1$ and $\mathbf{L}_2$ be two propositional logical systems presented as Hilbert style systems, using the set of connectives $\mathbf{C}_1$ and $\mathbf{C}_2$ respectively and based on the atomic propositions Q_1 and Q_2 respectively.

To have some specific examples, take $\mathbf{L}_1$ as intuitionistic implication "$\to$" based on Q_1 and let $\mathbf{L}_2$ be the same "$\to$" based on Q_2 and let $Q = Q_1 \cap Q_2$. We chose to base $\mathbf{L}_1$ and $\mathbf{L}_2$ on the same connective "$\to$" but we could have chosen $\mathbf{L}_1$ based on "$\to$" and say $\mathbf{L}_2$ based on "$\to$," giving "$\to$" the Hilbert axioms of strict **S4** implication.

It is important to note the Hilbert proof theory of such a logic $\mathbf{L}$. We define

$\vdash^0_{\mathbf{L}} A$ if A is a substitution instance of an axiom. (Sub$\mathbf{L}$)

If $\vdash^m_{\mathbf{L}} A$ and $\vdash^n_{\mathbf{L}} A \to B$ then $\vdash^k_{\mathbf{L}} B$, $k = \max(m, n)$.[1] (MP$\mathbf{L}$)

The important point to note is that in (Sub$\mathbf{L}$) above the "substitution instance" is in the language of the logic $\mathbf{L}$.

So if we have two versions $\mathbf{L}_1$ and $\mathbf{L}_2$ of intuitionistic implication "$\to$," one based on atoms Q_1 and one based on atoms Q_2, then in (Sub$\mathbf{L}_i$) the

[1]The modus ponens part can differ from logic to logic. This variation is essential and reflects the rich variety of substructural logics. What is important for us is the Hilbert style generation of the theorems (as opposed to, for example, natural deduction elimination/introduction rules) and the next point we make about (Sub$\mathbf{L}$).

substitution is in the langauge based on $Q_i, i = 1, 2$.

Let us now consider the logic based on $Q_1 \cup Q_2$, call it $\mathbf{L}_{1,2}$. We can have two versions of substitution for axioms:

Full version. (Sub$\mathbf{L}_{1,2}$) = all substitution instances of axioms with formulas based on $Q_1 \cup Q_2$.

Restricted version. Take the union of the sets of all substitution instances of the axioms of formulas based on Q_1 and separately the set based on Q_2. In symbols we can say that we are taking (Sub$\mathbf{L}_1$) $\cup$ (Sub$\mathbf{L}_2$).

Clearly (Sub$\mathbf{L}_1$) $\cup$ (Sub$\mathbf{L}_2$) $\subseteq$ (Sub$\mathbf{L}_{1,2}$).

Let $\vdash_{\mathbf{L}1,2}$ be the logic based on $Q_1 \cup Q_2$ generated from (Sub$\mathbf{L}_{1,2}$), and let $\vdash_{\mathbf{L}_1+\mathbf{L}_2}$ be the logic generated from (Sub$\mathbf{L}_1$) $\cup$ (Sub$\mathbf{L}_2$) using (MP$\mathbf{L}_1$) and (MP$\mathbf{L}_2$) and the new rule of transitivity

If $\vdash A \to B$ and $\vdash B \to C$ then $\vdash A \to C$. (TR$\mathbf{L}_1\mathbf{L}_2$).

We need this rule because modus ponens applies in each language only and TR gives the connection between languages.

The following holds in some cases (including the case of intuitionistic logic).

Claim. (1) *and* (2) *are equivalent*:

(1) $\mathbf{L}_{1,2}$ *has interpolation,*

(2) $\vdash_{\mathbf{L}_1+\mathbf{L}_2} = \vdash_{\mathbf{L}_{1,2}}$.

To persuade ourselves of this claim, assume $\mathbf{L}_{1,2}$ has interpolation and assume $\vdash_{\mathbf{L}_{1,2}} A(p, q) \to B(q, r)$. (We take this example as a typical case.) Then we need to show that $\vdash_{\mathbf{L}_1+\mathbf{L}_2} A(p, q) \to B(q, r)$, namely that it can be generated from (Sub$\mathbf{L}_1$) $\cup$ (Sub$\mathbf{L}_2$). By interpolation there is an $H(q)$ such that

$$\vdash_{\mathbf{L}_{1,2}} A \to H \text{ and } \vdash_{\mathbf{L}_{1,2}} H \to B.$$

Since it is easy to show that $\mathbf{L}_{1,2}$ is a conservative extension of $\mathbf{L}_i$, we can now show that $\vdash_{\mathbf{L}_1+\mathbf{L}_2} A \to B$.

Now assume (2) and we show interpolation. This is actually simple. If we write $\vdash_1, \vdash_2$ for provability in languages 1 and 2 respectively, then $\vdash_{1+2}$ is the transitive closure of $\vdash_1 \cup \vdash_2$, while $\vdash_{1,2}$ is direct provability in the language of $Q_1 \cup Q_2$. So if $A \vdash_{1+2} B$ and A is in language 1 and B in language 2, then there must be a sequence of wffs such that

$$A \vdash_1 A_1 \vdash_1 A_2 \ldots \vdash_1 A_k \vdash_2 B_1 \vdash \ldots \vdash B.$$

The junction $A_k \vdash_2 B_1$ is the first point in the sequence where there is a change in provability. Since we are starting with language 1, we have to start with $\vdash_1$ and when we change to $\vdash_2, A_k$ must be in the common language. Thus A_k is an interpolant.

The above argument is deceptively easy. This is because the hard work is to establish (2). The reader should note that condition (2) is not just a technical condition (equivalent to interpolation) which has no meaning of its own. Far from it, we are touching here on the area of combining logics and combining varieties etc and this is a deep subject (see [Gabbay, 1998; Gabbay *et al.*, 2003; Gabbay, 2008].

To highlight the distinction we made between $\vdash_{1,2}$ and $\vdash_{1+2}$ consider the following example.

EXAMPLE 1 (Logics with conjunction only). Let "$\wedge$" and "&" be two conjunctions. Write rules for each:

(& Introduction) $\dfrac{A, B}{A\&B}$,

(& Elimination) $\dfrac{A\&B}{A}$, $\dfrac{A\&B}{B}$,

($\wedge$ Introduction) $\dfrac{A, B}{A \wedge B}$,

($\wedge$ Elimination) $\dfrac{A \wedge B}{A}$, $\dfrac{A \wedge B}{B}$.

If we combine the language and combine the rules we get $(A \wedge B) \equiv (A\&B)$ as follows:

$$A\&B \vdash A$$
$$A\&B \vdash B$$
$$A, B \vdash A \wedge B$$
and hence
$$A\&B \vdash A \wedge B$$

However, if we allow the rules to be applied only in each logic take transitive closure, this will not make the two conjunctions equal.

Let us try and prove

$$(A \wedge B) \wedge C \vdash (A \wedge B)\&C.$$

we can get

$$(A \wedge B) \wedge C \vdash (A \wedge B),$$
$$(A \wedge B) \wedge C \vdash C,$$

but we cannot use

$$A \wedge B, C \vdash (A \wedge B) \& C$$

because it is not purely in one language!

We can continue

$$A \wedge B \vdash A\ ,$$
$$A \wedge B \vdash B$$

and then collect and get

$$A, B, C \vdash (A \& B) \& C$$

and so

$$(A \wedge B) \wedge C \vdash (A \& B) \& C$$

but not mixed!

The concept of restricted combination applies to any form of proof system and not necessarily only to a Hilbert formulation. It is a version of the concept of amalgamation (properly formulated) which may be equivalent to interpolation (properly and respectively formulated).

Amalgamation is a fascinating subject. See for example the papers of H. Neumann in the *American Journal of Mathematics*, 590–625 (1948) and 491–540 (1949).

In these papers, Figure 3 is studied for groups

$$\begin{tikzcd} & G & \\ G_1 \arrow[ur, "j_1 \cdots"] & & G_n \arrow[ul, "j_n"'] \\ & H \arrow[ul, "i_1 \cdots"] \arrow[ur, "i_n"'] & \end{tikzcd}$$

Figure 3.

where given that H is isomorphically embedded into $G_1, \ldots, G_n$ through $i_1, \ldots, i_m$. We want an amalgamated group G containing $G_1, \ldots, G_n$ such that all the copies $i_1(H), \ldots, i_n(H)$ are the *same copy*.

Interpolation: View 9 — Definability

To see this point of view recall the implicit function theorem from first year calculus.

Let $G(x,y)$ be a differentiable function with $G(x_0,y_0)=0$ and $\dfrac{\partial G}{\partial y} \neq 0$ at (x_0,y_0).

Then there exists a unique function $y = h(x)$, defined near $x = x_0$ such that $h(x_0) = y_0$ and $G(x,h(x)) \equiv 0$ (identical 0). The function h is differentiable and

$$h'(x) = \frac{\partial G/\partial x}{\partial G/\partial y}\,.$$

The above is a very fundamental basic theorem which is used in proving many key theorems from differential equations to differential geometry and manifolds. It essentially says that under reasonable conditions if $G(x,y)=0$ defines y as a unique function of x implicitly, with $G(x_0,y_0)=0$, then an explicit solution/definition exists $y = h(x)$ in the vicinity of (x_0,y_0) (i.e., $y_0 = h(x_0)$.)

Let us look now at Beth's definability theorem in logic. Suppose we have (the exactly formulation of the Beth property can also depend on the logic):

$$A(p,q_1)\&A(p,q_2) \vdash q_1 \equiv q_2.$$

Then there exists a formula $h(p)$ such that

$$A(p,q) \vdash q \equiv h(p).$$

Interpolation is sometimes equivalent to Beth definability and sometimes not. It depends on the logic. Let us consider classical propositional logic.

Assume we have interpolation and we show Beth definability.
Assume that $A(p,q_1)\&A(p,q_2) \vdash q_1 \equiv q_2$.
Then $A(p,q_1)\&q_1 \vdash A(p,q_2) \to q_2$.
So for some $h(p)$

(a) $A(p,q_1) \vdash q_1 \to h(p)$,

(b) $h(p)\&A(p,q_2) \vdash q_2$.

Hence $A(p,q) \vdash q \equiv h(p)$.

The converse can also be proved in the classical logic. But it does not hold in intermediate logics.

Interpolation: View 10 — Interpolation by translation

This is actually a general methodology for solving problems in logic. Suppose we are given two logical systems $\vdash_1$ and $\vdash_2$ and $\vdash_2$ is well known to us (for example, it is classical or intuitionistic predicate logic). If we have a faithful translation τ from $\vdash_1$ to $\vdash_2$ then we can learn properties of $\vdash_1$ by translating into $\vdash_2$ and then translating back.

The general schema is as follows:

We have a theory Θ in $\vdash_2$ such that for any A, B of $\vdash_1$ we have

$$A \vdash_1 B \text{ iff } \Theta \cup \{\tau(A)\} \vdash_2 \tau(B) \qquad (*)$$

Let $A \vdash_1 B$ be an interpolation problem in $\vdash_1$ and assume Θ is in the common language and assume further that $\vdash_2$ does allow for interpolation. Then we have

$$\Theta, \tau(A) \vdash_2 \tau(B).$$

Hence there exists an interpolant I in $\vdash_2$ such that

(1) $\Theta, \tau(A) \vdash_2 I$, (2) $I \vdash_2 \tau(B)$.

If we can find an H in $\vdash_1$ such that $\tau(H) = I$, (i.e., if we can translate I back into $\vdash_1$) then

$$A \vdash_1 H \vdash_1 B.$$

Thus the problem of interpolation for $\vdash_1$ is reduced to the problem of translation back from $\vdash_2$ to $\vdash_1$, which is again the problem of expressive power in $\vdash_1$.

We have a chapter devoted to this method in the book [Gabbay and Maksimova, 2005].

By the way, the "logic by translation" is a general methodology. We can do "revision by translation" or any other property by translation.

Another aspect of translation is to translate a schematic family of logics L_i^1 into a respective family L_i^2 in a systematic way and see what happens. For instance, the translation of superintuitionistic logics into extensions of S4 was exploited to solve interpolation problem over S4.

There are other useful translations such as the translation of nonmonotonic systems into temporal logic, which can help with interpolation.

Interpolation: View 11 — Traditional studies

There are traditional questions a logician asks in any logic discipline and our subject has its share of this. The following are some such questions.

(1) Complexity of (the algorithms for finding) the interpolants.

(2) Given a logic L for which there is (or there is not) an interpolation theorem, what can we learn about neighbouring systems?

(3) Variations of the interpolation theorem and how they relate to each other.

(4) Same as 3 for concepts related to interpolation.

Remarks. We conclude this section with an observation. The reader should note that the views mentioned above are not just methods for proving interpolation which are conceptually subordinate to the concept of interpolation itself. These are methodologies of independent interest and independent ideas and in many cases these are independent communities of researchers which are active in their area independently of interpolation.

1.2 Interpolation in general logics

1.2.1 Historical background

In recent years, applied disciplines such as computer science, artificial intelligence and computational linguistics were under increasing commercial pressure to provide devices which help and/or replace the human in his daily activity. This pressure required the use of logic in the modelling of human activity and organisation on the one hand and to provide the theoretical basis for the computer program constructs on the other.

The increased demand for logic from computer science and artificial intelligence and computational linguistics accelerated the development of the subject. It pushed research forward, stimulated by the needs of applications. New logic areas became established and old areas were enriched and expanded.

The table at the end of the editorial to this Volume, will give our readers an idea of the landscape of logic and its relation to computer science and formal language and artificial intelligence.

Perhaps the most impressive achievement of philosophical logic as arising in the past decade has been the effective negotiation of research partnerships with fallacy theory. These subjects are becoming more and more useful in agent theory and intelligent and reactive databases.

1.2.2 General logics and interpolation

The discussion in the previous subsection has shown that the notion of what is a logical system has evolved considerably in recent years. Similarly the notion of what we mean by interpolation in such logics needs to be modified. This gives us the opportunity not only to address interpolation for new logics but also to sharpen our understanding of interpolation in the old traditional logics, as we develop new points of view during our dealings with the new logics.

The purpose of this subsection is to discuss all these new developments.

Our method of exposition is to isolate and discuss main features arising in the new logic which seem to cause us difficulty in formulating and solving the interpolation problem.

Challenge 1: Structured databases. We begin with classical propositional logic. This logic has a notion of a formula A (built up using atoms

Logic / Notion	Intuitionistic or classical logic	Modal logic	Lambek logic
Declarative unit (formula A)	formula	formula	formula
Theory (database) Δ	set of formulas	sequence of formulas	sequence of formulas
Input into a database	Input an element into a set	Append an element at the end of a sequence	Append an element into end of a sequence
Deletion from database	Delete from a set	Drop last element in a sequence	Drop last element in a sequence
Substitution of a database into another (cut)	$\Delta(x/\Gamma) = (\Delta - \{x\}) \cup \Gamma$	Take the first x out of the sequence Δ and replace by the sequence Γ	Take any element x out of Δ and replace by Γ.
Language	Classical implication	Strict implication	Lambek resource implication

Figure 4.

and the traditional connectives), a notion of a theory (set of formulas), a notion of how to add an assumption to a theory and how to take an assumption out of a theory, as well as a notion of substitution of one theory into another (as seen in formulation of a cut).

The new logics differ from classical and intuitionistic logic in that they need variations on the above notions.

Let us summarise these notions in a table and indicate how two other sample logics have different notions. We simplify our language and look only at pure implicational logics. Figure 4 summarises the results for classical/intuitionistic implication, strict implication and Lambek implication.

A database in intuitionistic logic is a set of formulas. A database in strict implication or Lambek is a sequence of formulas. The consequence relation is always of the form $\Delta \vdash A$, i.e.,

$$\text{Database } \vdash \text{formula.}$$

Thus classical and intuitionistic logic have the same syntax but different consequence relation $\vdash$. Strict implication and Lambek also have almost the same syntax to each other but different consequence relations. Our book

[Gabbay and Maksimova, 2005, Section 17.1] contains more comparison and details.

The cut theorem (if at all true) in the case of sequence database looks like the following:

Cut for strict K implication

$(A_1, \ldots, A_n) \vdash B$
and
$(C_1, \ldots, C_m) \vdash A_1$
imply
$(C_1, \ldots, C_m, A_2, \ldots, A_n) \vdash B$

Cut for Lambek implication

$(A_1, \ldots, A_n, X, D_1, \ldots, D_k) \vdash B$
and
$(C_1, \ldots, C_m) \vdash X$
imply
$(A_1, \ldots, A_n, C_1, \ldots, C_m, D_1, \ldots, D_k) \vdash B$

To give an idea what $\vdash$ means in the case of strict implication, we use the semantic interpretation in Kripke models:

- $(A_1, \ldots, A_n) \vdash B$ iff in every Kripke model and every sequence of worlds $(t_1, \ldots, t_n)$ such that $t_1Rt_2, t_2Rt_3, \ldots t_{n-1}Rt_n$ and such that $t_1 \vDash A_1$ and $t_2 \vDash A_2, \ldots,$ and $t_n \vDash A_n$, we also have that $t_n \vDash B$.

Figure 5 shows the situation:

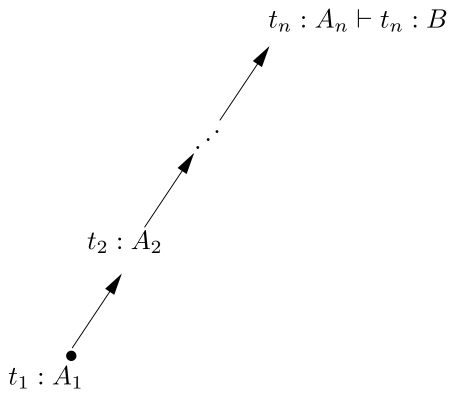


Figure 5.

The Lambek calculus implication requires semigroup (with operation $*$) semantics. We have

$(A_1, \ldots, A_n) \vdash B$ iff in every semigroup model and any $t_1, \ldots, t_n$ such that $t_1 \vDash A_1, \ldots, t_n \vDash A_n$ we have that $((t_1 * t_2) * \ldots) * t_n \vDash B$.

The additional difficulties arising for the problem of interpolation for logics where the database has additional structure is the challenge of the very formulation of the interpolation theorem.

We could of course write

$$A(p, q) \vdash B(q, r)$$

and ask for $H(q)$ such that $A \vdash H \vdash B$.

Since any logic of any complex structure will have to accept any single formula as a legitimate database, the above is always meaningful.

However, this setting may not be general enough to be useful (besides being a coward's way out!). We want a formulation which reflects the expressive properties of the logic and closed under inductive manoeuvres. If the logic has axioms or proof rules then $A \vdash B$ might be forced to reduce to $\Delta(A, B_i) \vdash C$ where B_i and C are subformulas and Δ is a structured database. In fact, this may be the very method for finding what structures are needed!

It makes more sense to try something like the following (for the case of the data being sequences)

$$A_1(p, q), B_1(q, r), A_2(p, q), \ldots, \vdash B$$

i.e., we want the languages to alternate. But then in this case we need to figure out first that the interpolant is probably going to be a sequence and what role is the interpolant sequence $H_1, H_2, \ldots$ to play in the picture? See [Gabbay and Olivetti, 2000], where new notions of chain interpolation were introduced.

The reader at this stage needs to realise the following points

1. Applied logic forces us to consider structured databases Δ which contain many formulas arranged in a specified structure (lists, trees, generally labelled, etc). The consequence relation is much more complex for such cases.

2. We need to formulate reasonable interpolation theorems for such structures and check their validity.

3. Many of the methods may not work here and may need to be specially modified. For example, algebraic methods will have a problem, how are we going to do algebraic logic for lists? Other methods such as semantical methods and proof theoretical methods will become more dominant. We may learn more about the old logics (classical, intuitionistic, modal), owing to the possibility of them being formulated as special limiting cases of the structured ones.

Before we conclude this subsection, let us talk more about structures. The most general logic with structured databases is that of labelled deductive systems. These systems have an algebra $\mathcal{A}$ of labels ($\mathcal{A}$ having algebraic and relational signature) and the formulas are "structured" by possibly several algebras of the form $\mathcal{A}$, i.e., our database Δ have the form $\Delta = (\mathcal{A}, \mathbf{f})$, where $\mathbf{f}$ is a function $\mathbf{f} : \mathcal{A} \to$ set of wffs, and $\mathcal{A}$ is one of the algebras we can use. We write $\mathbf{f}$ visually as $\{t_i : A_i\}, t_i \in \mathcal{A}$ where $A_i = \mathbf{f}(t_i)$. The consequence relation must therefore be defined in the form $\{t_i : A_i\} \vdash s : B$, for $t_i, s \in \mathcal{A}$.

We now have to consider an interpolation problem in the form:

Let $\mathbf{f} = \mathbf{f}_1 \cup \mathbf{f}_2$, where Domain $\mathbf{f}_1 \cap$ Domain $\mathbf{f}_2 = \varnothing$. Assume $\mathbf{f}_1(t_i)$ is in language 1 and $\mathbf{f}_2(t_j)$ is in language 2. Assume $\{t : \mathbf{f}(t)\} \vdash s : B$ with B in language 2. Can we eliminate $\mathbf{f}_1$ in some form and substitute by $\mathbf{f}_1'$ giving values in the common language?

Currently we have neither a general formulation nor a solution to this problem (see [Gabbay, 1996]).

Challenge 2: Nonmonotonicity. The interpolation challenges we had with strict implication arose because our database was structured. It was a list. But our logic was monotonic. Namely, if $\Delta \vdash A$ and Δ' *has more information than* Δ (this notion has to be defined) then also $\Delta' \vdash A$.

This condition is called *monotonicity*. Both strict implication and the Lambek implication are essentially monotonic.

In the case of strict implication, we can say that

a sequence $(A_1, \ldots, A_n, B_1, \ldots, B_k)$
has *more information than* $(D_1, \ldots, D_k)$
if for each $1 \leqslant i \leqslant k$ we have $(A_1, \ldots, A_n, B_1, \ldots, B_i) \vdash D_i$.

The Lambek implication is not technically monotonic but it is monotonic in spirit. If we change the definition of $\vdash$ a bit it will become monotonic.[2]

In a general labelled deductive system, we have to define the notion of "Δ has more information than Δ" using some homomorphic mapping of Δ' on Δ satisfying some provability conditions.[3]

We now address the problems arising from real nonmonotonicity.

EXAMPLE 2 (Negation as failure). Imagine that we go to a travel agent and ask for a flight from Novosibirsk to London *direct*. The agent looks up his list of flights and says that there is *no* such flight. All flights must connect through Moscow. We continue and ask whether there is a flight where the connection is such that one need not get off the plane (i.e., the plane just picks up more passengers at Moscow). Again the agent checks

---

[2]Lambek logic arose in natural language processing. A sentence like John is tall needs to be parsed and so it is present as np (representing John) and $np \to sentence$ (representing Tall) and the sentence is well formed if np, $np \to sentence \vdash sentence$. Thus John is Tall Tall should not be successful. But this is a technical nonmonotonicity.

[3]This, by the way, has not been published yet. It is in Volume 2 of Gabbay's Labelled Deductive Systems book.

his tables of flights and says that although there are no details about the connecting flight, the time gap is 4 hours and so he cannot believe it is the same plane (collecting more passengers).

The agent used *negation as failure*, a fundamental nonmonotonic principle, to deduce that there is no direct flight. The reasoning principle was that since such a flight was not listed, then there was no such flight!

This means that on a more formal basis we can do a deduction of the form

$$a, a \to q, \neg s \& q \to r \vdash r \tag{1}$$

The reason being that we can deduce $\neg s$ because s is atomic and is not listed in the database.

In fact, Horn clause logic programming with negation as failure will indeed give us r as a consequence, as suggested in (1). Consider now the interpolation question

$$a, a \to q \vdash (\neg s \& q \to r) \to r \tag{2}$$

This is supposed to be equivalent by the deduction theorem, if it holds, to the previous problem (1).

We now ask what is the common language? Is s in the common language? If the database says *no* to anything (atomic!) not listed in it, then its language is everything! It proves $\neg s \& \neg r \& \neg x \ldots$ etc, etc, all the atoms not listed in it. Indeed in the old days, when nonmonotonic logic was just starting to spread, many logicians insisted that there was no such thing as nonmonotonicity. They strongly claimed that if we are given some database Δ, then there are some principles (such as negation by failure, default, circumscription, etc, etc) which say that the actual database given is not Δ but Δ', where Δ' is the augmentation of Δ obtained by using these principles (in our case it is $\{a, a \to q, \neg s, \neg r, \neg p_1, \neg p_2, \ldots\}$), and that from Δ' everything can be derived in a monotonic way. This view is not correct, but we still have the problem of defining interpolation for such logics, and trying to decide what is the common language. Another problem in the case of negation as failure is that $\vdash$ may not be defined between clauses and the interpolant may be a clause.

To get ourselves even more agitated let us look at the following example:

EXAMPLE 3 (Transitivity). We have

(1) $\neg s \to r \vdash r$,

(2) $r \vdash s \to r$,

but

(3) $\neg s \to r \nvdash s \to r$ because (3) is equivalent to (4),

(4) $(\neg s \to r), s \nvdash r$ even though (3*) holds

(3*) $\neg s \to r \vdash \neg s$.

The first problem is that we have no transitivity. The second problem is that although (3*) holds, the minute we put s into the database (forming (4)), $\vdash \neg s$ is retracted! (This is why the system is nonmonotonic.)

So can we define and check interpolation in such a system?

Our case study section will define the above logic precisely. Meanwhile let us give you a semantic definition of another system — a circumscription system.

EXAMPLE 4 (Circumscription). (1) A classical model for the atoms $Q = \{q_1, q_2, \ldots\}$ is any function h $h : Q \mapsto \{0, 1\}$.

(2) We write $h_1 \leqslant_Q h_2$ iff for all $q \in Q, h_1(q) \leqslant h_2(q)$.

(3) Let A be a wff of classical logic. We say $h \vDash A$ iff the value of $h(A)$ of A is 1 when computed according to classical truth tables.

(4) We say h is a minimal model of A (notation $h \vDash^Q_m A$) iff $h \vDash A$ and there does not exist $h' \lneqq_Q h$ such that $h' \vDash A$.

(5) Define $A \vdash^Q_m B$ for A, B containing atoms from Q) by $A \vdash^Q_m B$ iff for any minimal model h of A we have $h \vDash B$. In symbols

$$h \vDash^Q_m A \text{ implies } h \vDash B.$$

(6) For example, we have $p \vee q \vdash^{\{p,q\}}_m \neg p \vee \neg q$.

Challenge. Formulate a reasonable interpolation question for $\vdash^Q_m$ and check its validity!

The reader can see that we again have a problem of common language. When we write $A(p, q) \vdash^Q_m B(q, r)$ what are the minimal models? Are they minimal as functions on $Q = \{p, q\}$ or on $Q = \{p, q, r\}$? In the second case we have $A(p, q) \vdash^Q_m \neg r$. In the first case we don't know.

The above discussion concentrated on the question of language, but there is also the question of consequence. If nonmonotonic logics may not be transitive (i.e., we may have $X \vdash Y$ and $Y \vdash Z$ but $X \nvdash Z$) then how do we know if interpolation fails because of lack of expressivity or because of lack of transitivity?

Suppose we have $A(p, q) \vdash B(q, r)$. Look for all $h(q)$ such that $A(p, q) \vdash h(q)$. If $H = \{h(q)\} \nvdash B(q, r)$ is it because of lack of logical expressivity or is it because $\vdash$ is not transitive?

Fortunately for our interpolation problems, nonmonotonic logics do have some structure to them. Almost every nonmonotonic logic $\vdash_1$ can be obtained in some systematic functorial[4] semantic or syntactic way from a maximal *monotonic* logic $\vdash_2$, having the main connecting property:

$$A \vdash_2 B \text{ implies } A \vdash_1 B \tag{$*$}$$

[4]By "functorial" we mean a universal construction on a logic which does not use any specific properties of it.

and some other connections arising from the construction.

This allows us to propose the following interpolation theorem for $\vdash_1$ (it may or may not hold).

Interpolation property for nonmonotonic consequence. Let $\vdash_1$ be a nonmonotonic consequence and let $\vdash_2$ be a monotonic consequence from which $\vdash_1$ is constructed. Then by *interpolation* we mean the following:

• If $A(p, q) \vdash_1 B(q, r)$ then there exist h_A and h_B in the common language such that $A \vdash_2 h_A \vdash_1 h_B \vdash_2 B$ (i.e., A monotonically interpolates to h_A which nonmonotonically proves h_B which then monotonically proves B).

The above overcomes the problem of lack of transitivity!

Challenge 3: Refined Interpolation. Our interpolation considerations so far were relative to the consequence relation. We wrote $A(p, q) \vdash B(q, r)$ and asked for an interpolant H. We did not ask how $\vdash$ works, that is we did not ask how it is presented to us (semantically, proof theoretically, as a Hilbert system, via a translation into another logic, etc) and then asked for the interpolant to be presented accordingly in some specific way. We were quite happy with any H, satisfying only the "$\vdash$" interpolating requirements. There are, however, several compelling reasons for us to insist on more from H. We want H to be influenced/tailored/be more meaningful in the context of the way the logic is presented. There are strong arguments in favour of this approach.

1. The view that a logic is just a consequence relation or just the set of its theorems is being seriously challenged by the emergence of the new logics. There are strong proof theoretical and there are enormously successful semantic methodologies that run across logics. We have tableaux methods, Gentzen methods, resolution methods, goal directed methods, truth table matrix methods, Hilbert methods, possible world semantics, algebraic semantics, etc. All these methodologies are cognitively distinct and easily identifiable and each can characterise a rich variety of diverse logics by executing minor changes in its procedures. Take for example three logics:

(a) classical logic,

(b) intuitionistic logic,

(c) Łukasiewicz n-valued logic.

Presented through the truth table matrix method, classical logic and Łukasiewicz logic are "brother and sister." Presented through Gentzen formulation, classical logic and intuitionistic logic are very close while Łukasiewicz logic is a pain in the neck! Intuitionistic logic has no truth presentation. It seems that the landscape of logics is a two dimensional grid (see Figure 1.6).

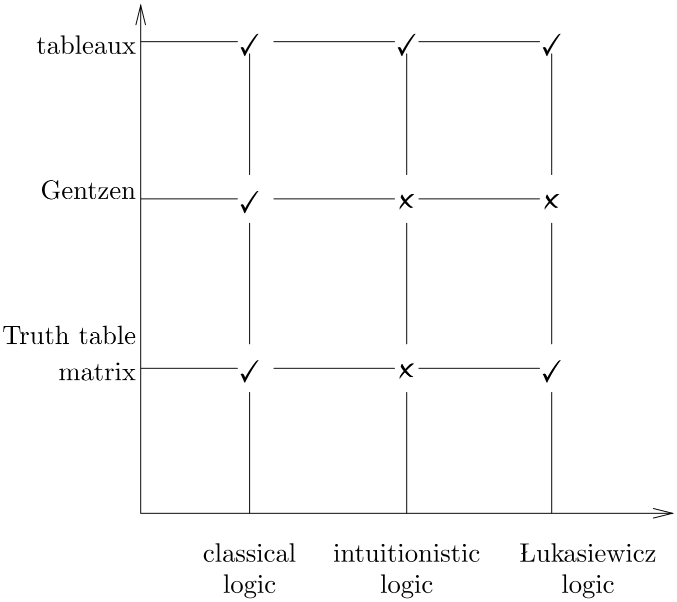


Figure 1.6

So classical propositional logic Gentzen presentation is *not* the same logic as classical propositional logic truth table presentation. So we may expect our interpolation demands to reflect that.

2. The connection of interpolation with computer science requires us to interpolate on the proof procedures which correspond to the programming steps. The set of theorems (which defines the logic) is meaningless in this context. This means that we should be more interested in finding interpolants relative to proofs which are in a sense *uniform across logics*, given a fixed database. So to be absolutely clear about this notion, given Δ and a family of logics L_1, L_2, all using variations of the same proof procedure π_i (π_i is a variation of some generic π) and given a common sublanguage $Q \subseteq Q_\Delta$ we find an interpolant $\mathbf{H}$ which is a sort of functional projection, so that if π proves B from Δ then $\mathbf{H}(\pi_i)$ proves B from Δ through some H_i, which can be identified in $\mathbf{H}(\pi_i)$.

Challenge 4: Interpolation and Abduction. Abduction is a general process for finding missing premisses. Thus if $\Delta \nvdash q$ we are looking for $\mathbb{A}$ such that $\Delta, \mathbb{A} \vdash q$. Processes that look for possible such $\mathbb{A}$'s and reason about them are abductive processes. This is a very central area in common sense reasoning and artificial intelligence as well as in philosophy of science (new scientific theories can be found by abduction/explanation). See the book [Gabbay and Woods, 2004]. For a quick lesson in abduction consider the following abduction problem:

$$p \to q \vdash ?q.$$

We want to know what data $\mathbb{A}$ to add to $\{p \to q\}$ to make it prove q. There are three typical options for abduction for the case of $\Delta \vdash ?q$.

Option 1. Simply add q (i.e., $\mathbb{A} = \{q\}$). This ia an easy way out and we are adding the maximal logical content. So in our case the new database becomes $\{p \to q, q\}$.

Option 2. Add $\Delta \to q$. This is a minimal approach. So in our case the new database becoems $\{p \to q, (p \to q) \to q\}$.

Option 3. This is an intermediate option and it depends on the computation procedure. We simply follow the computation (which fails!) and help it along with additional data whenever it gets stuck. In our case the abduced wff is $\{p\}$ because the computation uses modus ponens. So the new database is $\{p, p \to q\}$

We now ask: how is abudction connected with interpolation?

Consider the interpolation problem

$$\Delta_1, \Delta_2 \vdash q_2.$$

Assume that q_2 is genuinely in the language of Δ_2. The interpolation problem requries an $\mathbb{H}$ in the common language such that

$$\Delta_1 \vdash \mathbb{H} \text{ and } \mathbb{H}, \Delta_2 \vdash q_2.$$

Write this situation as

$$\Delta_1 \vdash \mathbb{H} \vdash \Delta_2 \to q_2.$$

Now consider the abduction problem $\Delta_2 \vdash ?q_2$. Clearly we need to abduce some $\mathbb{A}$ such that $\Delta_2, \mathbb{A} \vdash q_2$. Suppose we can fine-tune the abduction algorithm to find an $\mathbb{A}$ in the common language. Call it $\mathbb{H}$. If we also have $\Delta_1 \vdash \mathbb{H}$ then we found our interpolant. So we need a mechanism that can *modify* an abduction algorithm for $\Delta_2 \vdash ?q_2$ which can make use of the fact that $\Delta_1 \vdash \Delta_2 \to q_2$ and try and abduce in the common language. So our interpolation problem becomes a *modified abudction* problem.

By way of illustration, let us analyse the following database:

$$\Delta_1 = \{a \to x, \{x, r\} \to s\},$$
$$\Delta_2 = \{z, z \to r, (a \to s) \to p\}.$$

Here the common language is $\{a, r, s\}$.

Clearly $\Delta_1, \Delta_2 \vdash p$.

Let us try to prove p from Δ_2 only. What do we need to add to Δ_2 to make it prove p?

1. We can add p itself.

2. We can add $\Delta_2 \to p$.

Clearly (1) and (2) are two extreme cases (strongest and weakest assumptins). The following (3) is intermediate, as we disucssed.

3. Follow the computation and help it by adding the necessary assumptions. This option is good for interpolation because it allows us to

(a) restrict our additions to the common language and

(b) look at Δ_1 for clues of what to add.

So let us do it. Since we have in Δ_2 the clause $(a \to s) \to p$, we continue the computation by trying to prove $a \to s$. We therefore add a to the data and try and prove s. Nothing in the data can give us s but we do have r. So let us abduce and try to add $r \to s$, but since this is done under the assumption a, we need to add to Δ_2 the clause $a \to (r \to s)$.

Our candidate to add is then

$$\mathbb{A} = \{a \to (r \to s)\}.$$

But we also want $\Delta_1 \vdash \mathbb{A}$, so we want to make $\mathbb{A}$ as weak as possible. However in this case we do have $\Delta_1 \vdash a \to (r \to s)$.

2 INTERPOLATION AND DEFINABILITY IN MODAL AND INTUITIONISTIC LOGICS

In this section we present the basic results on interpolation in modal and intuitionistic logics, and also in some related logics.

We study interrelations of definability, interpolation and joint consistency, and present a proof of Lyndon's interpolation property (LIP) for some modal logics, in particular, for quantified K, K4 and S4, and for some other predicate systems. Also the propositional S5 has LIP. At the same time, the quantified extension of S5, as well as other systems satisfying the Barcan formula, have neither Lyndon's nor Craig's interpolation, nor Beth's property [Fine, 1979]. There are normal modal logics which have CIP but do not possess LIP. Also Craig's interpolation property holds for the most known propositional modal logics, in particular, for Grzegorczyk's logic Grz and for the provability logic G.

In Section 2.3 we give a brief overview of the basic results on interpolation. The detailed study can be found in [Gabbay and Maksimova, 2005].

First of all we consider

2.1 Interrelations between interpolation, definabillity and joint consistency

A *normal modal logic* is any set of modal formulas containing all the classical tautologies and the axioms $\Box(A \to B) \to (\Box A \to \Box B)$, $\Diamond A \equiv \neg\Box\neg A$ and closed under the rules R1: $A, (A \to B)/B$ (modus ponens), R2: $A/\Box A$ (necessitation) and substitution. The minimal normal modal logic is denoted by K. The family of all normal modal logics, which are extensions of a modal logic L, is denoted by $NE(L)$. The least normal extension of L containing a set Ax of formulas is denoted by $L + Ax$.

We use standard denotations for some members of $NE(\mathrm{K})$:

$$\begin{aligned}
&\mathrm{D} = \mathrm{K} + \Diamond\top,\\
&\mathrm{K4} = \mathrm{K} + (\Box p \to \Box\Box p),\\
&\mathrm{D4} = \mathrm{D} + (\Box p \to \Box\Box p),\\
&\mathrm{T} = \mathrm{K} + (\Box p \to p),\\
&\mathrm{S4} = \mathrm{K4} + (\Box p \to p),\\
&\mathrm{G} = \mathrm{K4} + (\Box(\Box p \to p) \to \Box p),\\
&\mathrm{Grz} = \mathrm{S4} + (\Box(\Box(p \to \Box p) \to p) \to p),\\
&\mathrm{Grz.2} = \mathrm{Grz} + (\Diamond\Box p \to \Box\Diamond p),\\
&\mathrm{K4.3} = \mathrm{K4} + (\Box(p\&\Box p \to q) \vee \Box(q\&\Box q \to p),\\
&\mathrm{S4.3} = \mathrm{S4} + \mathrm{K4.3},\\
&\mathrm{S5} = \mathrm{S4} + (p \to \Box\Diamond p).
\end{aligned}$$

If L is a normal modal logic, its natural quantified extension is denoted by LQ. Also we consider quantified modal logics with the *Barcan formula:* $\forall x\Box A \to \Box\forall x A$.

The relational semantics of the first-order modal language is defined as follows. Let W be a set and R be a binary relation of accessibility on W. Let D be some non-empty set (a domain). A *first-order Kripke frame over* D is a triple

$$\mathbf{W} = \langle W, R, \{D_w\}_{w\in W}\rangle,$$

where for any w, D_w is a non-empty subset of D (called by *domain of* w) and also the *monotonicity condition* is satisfied:

$$\text{for all } w, w' \in W,\ wRw' \text{ implies } D_w \subseteq D_{w'}.$$

The frame $\mathbf{W} = \langle W, R, \{D_w\}_{w\in W}\rangle$ is a *constant domain frame,* whenever $D_w = D_{w'} = D$ for all w, w' in W.

A *model* is a frame together with truth-relation. If L is a logic, a model $\mathbf{M}$ is said to be an *L-model* if all axioms of L are true at all worlds in W; a frame $\mathbf{W}$ is an *L-frame* whenever all axioms of L are valid in all models based on $\mathbf{W}$. It is well known that the logic K is characterized by all frames, D by serial frames, K4 by transitive frames, T by reflexive frames, S4 by reflexive and transitive frames, S5 by frames with the total relation R. The Barcan formula is valid in a frame $\mathbf{W}$ if and only if $\mathbf{W}$ is a constant domain frame.

Interpolation and definability notions play an important part in the mathematical logic. We consider various versions of these properties. We assume that first order language contains the propositional constant $\top$.

The Beth definability theorem [Beth, 1953] states that any predicate implicitly definable in a first order theory is explicitly definable.

Let P be a list of predicate letters and individual constants and variables. We write $A(P)$ to indicate that all predicates, constants and free variables of the formula A are contained in P.

The *Beth definability property* $B1$ can be formulated for the predicate logics as follows:

Let $A(P, X)$ be a first order formula, X and Y n-ary predicate variables outside of $P, \mathbf{u} = u_1, \ldots, u_n$ a list of individual variables outside of P. If

$$L \vdash A(P, X) \& A(P, Y) \to \forall \mathbf{u}(X(\mathbf{u}) \equiv Y(\mathbf{u}))$$

then there exists a formula $B(P, \mathbf{u})$ such that

$$L \vdash A(P, X) \to \forall \mathbf{u}(X(\mathbf{u}) \equiv B(P, \mathbf{u})).$$

Here the first formula means that $A(P, X)$ *implicitly defines* X, and the conclusion states that $A(P, X)$ *explicitly defines* X; the formula $B(P, \mathbf{u})$ is said to be an *explicit definition* of X.

The Beth property B1 is closely connected with the Craig interpolation property.

The *Craig interpolation property (CIP)* of a logic L is the following:

Let P, Q, R be disjoint sets consisting of predicate symbols and individual constants. If

$$L \vdash A(P, Q) \to B(P, R)),$$

then there exists a formula $C(P)$ such that

$$L \vdash A(P, Q) \to C(P))$$

and

$$L \vdash C(P) \to B(P, R)).$$

Craig [Craig, 1957] proved that in the classical first-order logic the interpolation property implies the Beth property B1. In the same way one can prove

LEMMA 5. *Let L be any modal logic. If L has CIP then L possesses* B1.

We defined the Beth property in a syntactic way. Often other definitions are applied, where the implicit and/or explicit definability is given semantically. Assume K is a class of Kripke models. Then it determines a semantical consequence relation. For a set of sentences Γ and a sentence B we may say that Γ semantically implies B on the class K and write $\vDash_K \Gamma \to B$ if for any model $\mathbf{M} = \langle W, R, \{D_w\}_{w \in W}, \vDash\rangle$ in K and for any $w \in W$ we have

$$w \vDash \Gamma \Rightarrow w \vDash B.$$

Let $A(P, X)$ be a first order formula, X and Y n-ary predicate variables outside of $P, \mathbf{u} = u_1, \ldots, u_n$ a list of individual variables outside of P, and let K be a class of models. We may say that X is implicitly definable by $\Gamma(P, X)$ on the class K if

$$\vDash_K \Gamma(P, X), \Gamma(P, Y) \to \forall \mathbf{u}(X(\mathbf{u}) \equiv Y(\mathbf{u}));$$

X is explicitly definable by $\Gamma(P, X)$ on the class K if there exists a formula $B(P, \mathbf{u})$ such that

$$\vDash_K \Gamma(P, X) \to \forall \mathbf{u}(X(\mathbf{u}) \equiv B(P, \mathbf{u})).$$

Then one can change the definition of the Beth property replacing its premise and/or its conclusion by the new definition of implicit or explicit definability. It depends on the class K whether the new definition will be equivalent to the old one or not. A sufficient condition for such an equivalence is strong completeness of the logic L with respect to the class K. In particular, the equivalence holds when L is one of the logics K, K4, S4, S5 and K is the class of models based on L-frames. This immediately follows from Strong Completeness Theorem (see Theorem 15(i) below).

Fix any predicate modal logic L. For any sets Γ and Δ of formulas, write $\Gamma \to_L \Delta$, if there exist formulae $A_1, \ldots, A_n$ in Γ and $B_1, \ldots, B_m$ in Δ such that $((A_1 \& \ldots \& A_n) \to (B_1 \vee \ldots B_m))$ is in L. A set Γ is L- *consistent* if there is no formula B such that $\Gamma \to_L B$ and $\Gamma \to_L \neg B$ for some formula B. We say that a set Γ of sentences of the language $\mathcal{F}$ is an *L-theory* of this language if Γ is closed with respect to $\to_L$, i.e., if $\Gamma \to_L B$ for some sentence B of $\mathcal{F}$ then $B \in \Gamma$. A set Γ of sentences of $\mathcal{F}$ is a *set of axioms* of an L-theory T if

$$T = \{B | B \text{ is a sentence of } \mathcal{F} \text{and } \Gamma \to_L B\}.$$

It is clear that a L-theory is L-consistent if and only if it does not contain $\bot$. A theory Γ of the language $\mathcal{F}$ is *complete* if it contains A or $\neg A$ for any sentence A of $\mathcal{F}$.

In classical modal logics the Craig interpolation property is equivalent to the following property RCP that is an analog of the Robinson Consistency Theorem:

(RCP) Let P, Q, R be disjoint lists of predicate symbols and individual constants, and let Γ be a L-consistent L-theory of the language $\mathcal{F}(P, Q)$, Δ a L-consistent L-theory of $\mathcal{F}(P, R)$ such that $\Gamma \cap \Delta$ is a complete and L-consistent L-theory of the language $\mathcal{F}(P)$. Then the set $\Gamma \cup \Delta$ is L-consistent.

In order to prove the equivalence between CIP and RCP, we need some definitions. Let L be a modal logic. Let P, Q, R be disjoint lists of predicate symbols and individual constants, $\mathcal{L}_1 = \mathcal{F}(P, Q)$, $\mathcal{L}_2 = \mathcal{F}(P, R)$. We restrict ourselves with considering at most countable languages. A pair of sets (T_1, T_2), where $T_i \subseteq \mathcal{L}_i$ for $i = 1, 2$, is *L-separable* if there is a formula C in $\mathcal{L}_0 = \mathcal{L}_1 \cap \mathcal{L}_2$ such that $T_1 \to_L C$ and $T_2 \to_L \neg C$; such C is called an *interpolant*. A pair is *L-inseparable* if it is not L-separable. An L-inseparable pair (T_1, T_2) is *$(L, \mathcal{L}_1, \mathcal{L}_2)$-complete* if T_i is $\mathcal{L}_i$-complete for $i = 1, 2$. Note that if (T_1, T_2) is $(L, \mathcal{L}_1, \mathcal{L}_2)$-complete and $A \in \mathcal{L}_1 \cap \mathcal{L}_2$, then $A \in T_1$ iff $A \in T_2$.

One can prove the following lemma for any modal logic L.

LEMMA 6. (i) *Every L-inseparable pair (T_1, T_2), where $T_i \subseteq \mathcal{L}_i$ for $i = 1, 2$, can be extended to a $(L, \mathcal{L}_1, \mathcal{L}_2)$-complete pair.*

(ii) *For each $(L, \mathcal{L}_1, \mathcal{L}_2)$-complete pair (T_1, T_2) and for every formula $A \in \mathcal{L}_i$,*

$$T_i \to_L A \iff A \in T_i.$$

In particular, $\perp \notin T_i$ and for all $A, B \in \mathcal{L}_i$,

$$L \vdash A \equiv B \Rightarrow (A \in T_i \iff B \in T_i).$$

THEOREM 7. *A modal logic L has* CIP *if and only if it has* RCP.

Proof. Let L have CIP. Let Γ be a L-consistent L-theory of the language $\mathcal{F}(P, Q)$, Δ a L-consistent L-theory of $\mathcal{F}(P, R)$, where P, Q, R be disjoint lists of predicate symbols and individual constants, and $\Gamma \cap \Delta$ is an L-complete L-theory of $\mathcal{F}(P)$. Suppose that $\Gamma \cup \Delta$ is L-inconsistent. Then $L \vdash A \& B \to \perp$, where A and B are conjunctions of some formulas in Γ and Δ respectively. Therefore, $L \vdash A \to \neg B$ and, by CIP, there exists a $C(P)$ such that $L \vdash A \to C(P)$ and $L \vdash C(P) \to \neg B$. Hence $C(P) \in \Gamma$ and $\neg C(P) \in \Delta$. If $C(P) \in \Gamma \cap \Delta$ then Δ is L-inconsistent. If $\neg C(P) \in \Gamma \cap \Delta$ then Γ is L-inconsistent. So $\Gamma \cap \Delta$ is incomplete. Thus L has RCP.

For the converse, assume that L has RCP. Let $L \vdash A(P, Q) \to B(P, R)$. Suppose that there is no interpolant for this formula. It follows that the pair $(\{A(P, Q)\}, \{\neg B(P, R)\})$ is L-inseparable. By Lemma 6 it can be extended to a $(L, \mathcal{L}_1, \mathcal{L}_2)$-complete pair (Γ, Δ). Then $\Gamma \cap \Delta$ is an $\mathcal{L}_0$-complete theory, and $\Gamma \cup \Delta$ is L-consistent by RCP. We note that $\Gamma \cup \Delta \to_L \neg(A(P, Q) \to B(P, R))$ contradicting the assumption. ■

2.2 Lyndon's interpolation in some modal systems

In this section we consider the Lyndon interpolation property (LIP) which is stronger then CIP. We prove that the most known quantified modal systems have this interpolation property, in particular, it holds for quantified logics K, K4 and S4 and for the propositional S5. Also some superintuitionistic logics have Lyndon's interpolation property. At the same time, there are normal extensions of the propositional S5 which have the Craig interpolation property and do not have the Lyndon interpolation. Some examples are given in [Gabbay and Maksimova, 2005]. Also the Craig interpolation property is known for the logics Grz and G. K. Fine [Fine, 1979] has proved that the quantified system S5 has neither interpolation nor the Beth property.

We use semantic methods in style of [Gabbay, 1972] to obtain interpolation theorems for quantified modal logics (see Theorem 2.10 below). The proofs were published in [Maksimova, 1982a]. Other proofs, also for nonnormal systems, can be found in [Fitting, 1983]. In this section we consider first order languages without equality and without functional symbols.

First we give some definitions.

Let us define notions of *negative* and *positive occurrences* of a subformula in a formula inductively as follows:

(i) occurrence of subformula A in A is positive,

(ii) occurrence of A in a positive (negative) occurrence of $(A\&B)$, $(B\&A)$, $(A \vee B)$, $(B \vee A)$, $(B \to A)$, $\Box A$, $\Diamond A$, $\exists xA$, $\forall xA$ is positive, (negative),

(iii) occurrence of A in a positive (negative) occurrence of $(A \to B), \neg A$ is negative (positive).

An occurrence of a symbol X in a formula A is called to be *positive (negative)*, if it is situated in a positive occurrence of some subformula of A.

For any set Γ of formulas, denote by $\Omega^+(\Gamma)$ the set of all predicate symbols occurring positively in formulae of Γ and by $\Omega^-(\Gamma)$ the set of all predicate symbols occurring negatively in formulae of $\Gamma, \Omega(\Gamma) = \Omega^+(\Gamma) \cup \Omega^-(\Gamma)$; $D(A)$ is the set of individual constants and free variables of a formula A.

The *Lyndon interpolation property (LIP)* in a logic L is the following:

If $(A \to B)$ is in L, then there exists a formula C such that $(A \to C)$ and $(C \to B)$ are both in L and also $D(C) \subseteq D(A) \cap D(B), \Omega^+(C) \subseteq \Omega^+(A) \cap \Omega^+(B)$ and $\Omega^-(C) \subseteq \Omega^-(A) \cap \Omega^-(B)$.

This formula C is called a *Lyndon interpolant of* $(A \to B)$.

Say, that a formula A is *reduced* if it is built from atomic formulas and their negations by use of $\&, \vee, \Box, \Diamond, \exists, \forall$.

LEMMA 8. *Any quantified modal formula A is equivalent in* K *to some reduced formula A', where $D(A') = D(A)$, $\Omega^+(A) = \Omega^+(A')$ and $\Omega^-(A) =$*

$\Omega^-(A')$.

For proof, one can use the replacement lemma and the following equivalences of K:

$$\neg\neg A \equiv A,\ (A \to B) \equiv (\neg A \vee B),\ \neg(A\&B) \equiv (\neg A \vee \neg B),$$
$$\neg(A \vee B) \equiv (\neg A\&\neg B),\ \neg\Box A \equiv \Diamond\neg A,\ \neg\Diamond A \equiv \Box\neg A,$$
$$\neg\forall x A \equiv \exists x\neg A,\ \neg\exists x A \equiv \forall x\neg A.$$

Let Γ be a set of formulas. Denote by $D(\Gamma)$ the set of all individual constants and free variables contained in Γ. Denote by $\mathcal{L}$ (sometimes with indices) a language consisting of reduced formulas, i.e., the set of all reduced formulas A satisfying the conditions $\Omega^+(A) \subseteq \Omega^+(\mathcal{L}), \Omega^-(A) \subseteq \Omega^-(\mathcal{L}), D(A) \subseteq D(\mathcal{L})$. If Γ is a set of reduced formulas, denote by $\mathcal{L}(\Gamma)$ the least language containing Γ. Languages $\mathcal{L}$ and $\mathcal{L}'$ are *correlated* if $D(\mathcal{L}') = D(\mathcal{L})$. A language $\mathcal{L}'$ is said to be an *inessential extension* of $\mathcal{L}$, whenever $D(\mathcal{L}') \subseteq D(\mathcal{L}), \Omega^+(\mathcal{L}') = \Omega^+(\mathcal{L}), \Omega^-(\mathcal{L}') = \Omega^-(\mathcal{L})$.

Fix any predicate modal logic L. For any sets Γ and Δ of formulas, write $\Gamma \to_L \Delta$, if there exist formulae $A_1, \ldots, A_n$ in Γ and $B_1, \ldots, B_m$ in Δ such that $((A_1\& \ldots \&A_n) \to (B_1 \vee \ldots B_m))$ is in L. By *a theory* of the language $\mathcal{L}$, we mean any set $T \subseteq \mathcal{L}$ satisfying the condition: $(A \in \mathcal{L}$ and $T \to_L A) \Rightarrow A \in T$. If $F \subseteq \mathcal{L}$ satisfies the dual condition: $(A \in \mathcal{L}$ and $A \to_L F) \Rightarrow A \in F$, then F is *a L-cotheory* of the language $\mathcal{L}$. Denote by $Th(\mathcal{L}, L)$ the set of all L-theories and by $CTh(\mathcal{L}, L)$ the set of all L-cotheories of the language $\mathcal{L}$.

Let $\mathcal{L}'$ and $\mathcal{L}''$ be correlated languages, $\mathcal{L} = \mathcal{L}' \cap \mathcal{L}'', \Gamma \subseteq \mathcal{L}'$ and $\Delta \subseteq \mathcal{L}''$. A pair (Γ, Δ) is called *L-separable in* $(\mathcal{L}', \mathcal{L}'')$, whenever there exists a formula A of the language $\mathcal{L}$ such that $\Gamma \to_L A$ and $A \to_L \Delta$. Note, that the pair (T, F), where $T \in Th(\mathcal{L}', L)$ and $F \in CTh(\mathcal{L}'', L)$, is L-inseparable in $(\mathcal{L}', \mathcal{L}'')$ if and only if $T \cap F = \emptyset$.

A pair (T, F) is called *L-saturated in* $(\mathcal{L}', \mathcal{L}'')$ if the following conditions are satisfied:

1) $T \in Th(\mathcal{L}', L)$ and $F \in CTh(\mathcal{L}'', L)$;
2) $T \cap F = \emptyset$, i.e., (T, F) is L-inseparable in $(\mathcal{L}', \mathcal{L}'')$;
3) $(A \vee B) \in T \Rightarrow (A \in T$ or $B \in T)$;
4) $\exists x A(x) \in T \Rightarrow (\exists c \in D(\mathcal{L}'))(A(c) \in T)$;
5) $(A\&B) \in F \Rightarrow (A \in F$ or $B \in F)$;
6) $\forall x A(x) \in F \Rightarrow (\exists c \in D(\mathcal{L}''))(A(c) \in F)$.

For fixed L, sometimes use terms "separable" and "saturated" instead of "L-separable" and "L-saturated."

LEMMA 9. *Let (Γ, Δ) be L-inseparable in $(\mathcal{L}', \mathcal{L}'')$. Then*

(i) *if $\Diamond A \in \Gamma$, then the pair $(\{A\} \cup \{A' | \Box A' \in \Gamma\}, \{B' | \Diamond B' \in \Delta\})$ is L-inseparable in $(\mathcal{L}', \mathcal{L}'')$;*

(ii) *if* $\Box B \in \Delta$*, then the pair* $(\{A' | \Box A' \in \Gamma\}, \{B\} \cup \{B' | \Diamond B' \in \Delta\})$ *is* L*-inseparable in* $(\mathcal{L}', \mathcal{L}'')$*.*

Proof. (i) Assume that $L \vdash ((A \& A_1 \& \ldots \& A_n) \to C)$ and $L \vdash (C \to (B_1 \vee \ldots \vee B_m))$ for some formula $C \in \mathcal{L} = \mathcal{L}' \cap \mathcal{L}''$, where $\Box A_1, \ldots, \Box A_n \in \Gamma$ and $\Diamond B_1, \ldots, \Diamond B_m \in \Delta$. Then $\Diamond C \in \mathcal{L}, L \vdash ((\Diamond A \& \Box A_1 \& \ldots \& \Box A_n \to \Diamond C$ and $L \vdash (\Diamond C \to (\Diamond B_1 \vee \ldots \vee \Diamond B_m))$, so the pair (Γ, Δ) is L-separable in $(\mathcal{L}', \mathcal{L}'')$.

(ii) Suppose that $L \vdash ((A_1 \& \ldots \& A_n) \to C)$ and $L \vdash (C \to B \vee B_1 \vee \ldots \vee B_m))$ for some $C \in \mathcal{L} = \mathcal{L}' \cap \mathcal{L}''$, where $\Box A_1, \ldots, A_n \in \Gamma$ and $\Diamond B_1, \ldots, \Diamond B_m \in \Delta$. Then $\Box C \in \mathcal{L}, L \vdash ((\Box A_1 \& \ldots \& \Box A_n) \to \Box C)$ and $L \vdash (\Box C \to (\Box B \vee \Diamond B_1 \vee \ldots \vee \Diamond B_m))$, so the pair (Γ, Δ) is L-separable in $(\mathcal{L}', \mathcal{L}'')$. ■

LEMMA 10. *Let* $\mathcal{L}'$ *and* $\mathcal{L}''$ *be countable. Then any pair* L*-inseparable in* $(\mathcal{L}', \mathcal{L}'')$ *can be extended to a pair which is* L*-saturated in* $(\mathcal{L}'^*, \mathcal{L}''^*)$*, where* $\mathcal{L}'^*$ $(\mathcal{L}''^*)$ *is an inessential extension of* $\mathcal{L}'$ $(\mathcal{L}'')$*.*

Proof. Let (Γ, Δ) be L-inseparable in $(\mathcal{L}', \mathcal{L}'')$. We extend $\mathcal{L}'$ and $\mathcal{L}''$ adding a countable set $\{c_0, c_1, \ldots\}$ of new individual constants and denote extensions by $\mathcal{L}'^*$ and $\mathcal{L}''^*$ respectively. Note that (Γ, Δ) is L-inseparable also in $(\mathcal{L}'^*, \mathcal{L}''^*)$. If there were a formula C in $\mathcal{L}^* = \mathcal{L}'^* \cap \mathcal{L}''^*$, such that $\Gamma \to_L C$ and $C \to_L \Delta$, one would construct a formula C' in $\mathcal{L} = \mathcal{L}' \cap \mathcal{L}''$ replacing the new constants by new individual variables and binding them by universal quantifiers, then C' would separate (Γ, Δ) in $(\mathcal{L}', \mathcal{L}'')$.

Now enumerate all the formulas of $(\mathcal{L}'^*, \mathcal{L}''^*)$:

$$\mathcal{L}'^* = \{A_1, A_1, \ldots\}, \mathcal{L}''^* = \{B_1, B_2, \ldots\}.$$

Let $T_0 = \Gamma, F_0 = \Delta$. Further, let $F_{2n+1} = F_{2n}$ and define T_{2n+1} as follows: $T_{2n+1} = T_{2n} \cup \{A_n\}$, whenever the pair $(T_{2n} \cup \{A_n\}, F_{2n})$ is L-inseparable in $(\mathcal{L}'^*, \mathcal{L}''^*)$ and A_n does not begin with existential quantifier;

$T_{2n+1} = T_{2n} \cup \{A_n, A_n'(c)\}$, whenever the pair $(T_{2n} \cup \{A_n\}, F_{2n})$ is L-inseparable in $(\mathcal{L}'^*, \mathcal{L}''^*)$ and $A_n = \exists x A_n'(x)$ and c is the former new constant, which has no occurrences in formulas of $(T_{2n} \cup \{A_n\}, F_{2n})$; $T_{2n+1} = T_{2n}$, whenever the pair $(T_{2n} \cup \{A_n\}, F_{2n})$ is L-separable in $(\mathcal{L}'^*, \mathcal{L}''^*)$.

At step $2n+2$ let $T_{2n+2} = T_{2n+1}$ and define F_{2n+2} as follows: $F_{2n+2} = F_{2n+1}$, whenever the pair $(T_{2n+1}, F_{2n+1} \cup \{B_n\})$ is L-separable in $(\mathcal{L}'^*, \mathcal{L}''^*)$; $F_{2n+2} = F_{2n+1} \cup \{B_n, B_n'(c)\}$, whenever $B_n = \forall x B_n'(x)$, the pair $(T_{2n+1}, F_{2n+1} \cup \{B_n\})$ is L-inseparable in $(\mathcal{L}'^*, \mathcal{L}''^*)$ and c is the former new constant, which has no occurrences in formulas of $(T_{2n+1}, F_{2n+1} \cup \{B_n\})$; $F_{2n+2} = F_{2n+1} \cup \{B_n\}$ in other case. Evidently, $T_0 \subseteq T_1 \subseteq T_2 \subseteq, \ldots,$ $F_0 \subseteq F_1 \subseteq F_2 \subseteq \ldots$.

In a standard way, one can proof by induction on n that each pair (T_n, F_n) is L-inseparable in $(\mathcal{L}'^*, \mathcal{L}''^*)$.

Now define

$$T = \{A \in \mathcal{L}'^{*} | \exists n(T_n \to_L A)\}, F = \{B \in \mathcal{L}''^{*} | \exists n(B \to_L F_n)\}$$

and prove that (T, F) is L-saturated in $(\mathcal{L}'^{*}, \mathcal{L}''^{*})$.

It is obvious that $T \in Th(\mathcal{L}'^{*}, L)$ and $F \in CTh(\mathcal{L}''^{*}, L)$ and also (T, F) is L-inseparable in $(\mathcal{L}'^{*}, \mathcal{L}''^{*})$. Further, the following holds:

(i) for any formula A in $\mathcal{L}'^{*} : A \in T$ iff $(T \cup \{A\}, F)$ is L-inseparable,

(ii) for any formula B in $\mathcal{L}''^{*} : B \in F$ iff $(T, F \cup \{B\})$ is L-inseparable.

Indeed, let $A = A_n$. Then inseparability of $(T \cup \{A\}, F)$ implies inseparability of $(T_{2n} \cup \{A\}, F_{2n})$ so $A \in T_{2n+1} \subseteq T$. On the other hand, from $A \in T$ there follows inseparability of $(T \cup \{A\}, F)$ since (T, F) is is inseparable. (ii) can be proved in similar way.

Suppose that $(A \vee B) \in T$, but $A \notin T$ and $B \notin T$. Then both $(T \cup \{A\}, F)$ and $(T \cup \{B\}, F)$ would be separable, i.e., $T, A \to_L C$ and $C \to_L F$ and also $T, B \to_L C'$ and $C \to_L F$ for some $C, C' \in \mathcal{L}^{*} = \mathcal{L}'^{*} \cap \mathcal{L}''^{*}$. Therefore,

$$T, A \vee B \to_L C \vee C' \text{and } C \vee C' \to_L F,\ C \vee C' \in \mathcal{L}^{*},$$

so the pair (T, F) would be separable in $(\mathcal{L}'^{*}, \mathcal{L}''^{*})$, but it is not the case. So, we have

$$(A \vee B) \in T \text{implies } (A \in T \text{or } B \in T).$$

In analogous way, one can prove that

$$(A \& B) \in F \text{implies } (A \in F \text{or } B \in F).$$

At last, let $\exists x A(x) \in T, \exists x A(x) = A_n$. Then the pair $(T_{2n} \cup \{A\}, F_{2n})$ is L-separable; by construction, $A(c_k) \in T_{2n+1}$ for some k and 4) of the definition of L-saturated theory is proved. 6) can be obtained by analogy. It completes the proof of Lemma 10. ■

We say that a pair (Γ, Δ) is *satisfiable in a model*

$$\mathbf{M} = \langle W, R, \{D_w\}_{w \in W}, \vDash \rangle$$

if there is a $w \in W$ such that $w \vDash A$ for any $A \in \Gamma$ and $w \vDash \neg B$ for any $B \in \Delta$; a pair (Γ, Δ) is *satisfiable in a frame* $\mathbf{W} = \langle W, R, \{D_w\}_{w \in W} \rangle$ if it is satisfiable in some model based on $\mathbf{W}$.

THEOREM 11. *Let L be any of quantified modal logics* KQ, DQ, TQ, K4Q, D4Q *and* S4Q. *Let $\mathcal{L}'$ and $\mathcal{L}''$ be correlated languages, $\mathcal{L} = \mathcal{L}' \cap \mathcal{L}'', \Gamma \subseteq \mathcal{L}', \Delta \subseteq \mathcal{L}''$ and a pair (Γ, Δ) be L-inseparable in $(\mathcal{L}', \mathcal{L}'')$. Then (Γ, Δ) is satisfiable in some L-frame.*

Proof. By Lemma 10, (Γ, Δ) can be extended to a pair (T_0, F_0) which is L-saturated in $(\mathcal{L}'^*, \mathcal{L}''^*)$, where $\mathcal{L}'^*$ and $\mathcal{L}''^*$ are inessential extensions of $\mathcal{L}'$ and $\mathcal{L}''$ respectively. Let

$$W_0 = \{(T_0, F_0)\}.$$

Suppose that W_k is already built and all its elements are L-saturated pairs in some suitable countable correlated languages. We use Lemmas 9 and 10 to extend W_k. Let $(T, F) \in W_k$ be any pair L-saturated in $(\mathcal{L}_1, \mathcal{L}_2)$. For each formula $\Diamond A \in T$ we extend W_k adding a pair (T', F') L-saturated in some suitable countable $(\mathcal{L}_1', \mathcal{L}_2')$, such that $T' \supseteq \{A\} \cup \{A' | \Box A' \in \Gamma\}, F' \supseteq \{B' | \Diamond B' \in \Delta\}, D(\mathcal{L}_1') = D(\mathcal{L}_2'), \supseteq D(\mathcal{L}_1) = D(\mathcal{L}_2), \mathcal{L}_i'$ is an inessential extension of $\mathcal{L}_i$. Also for each formula $\Box B \in F$ we add to W_k a pair (T', F') L-saturated in suitable countable $(\mathcal{L}_1', \mathcal{L}_2')$ such that $T' \supseteq \{A' | \Box A' \in \Gamma\}, F' \supseteq \{B\} \cup \{B' | \Diamond B' \in \Delta\}, D(\mathcal{L}_1') = D(\mathcal{L}_2') \supseteq D(\mathcal{L}_1) = D(\mathcal{L}_2), \mathcal{L}_i'$ is an inessential extension of $\mathcal{L}_i$. So W_{k+1} arises from W_k by adding new pairs for all (T, F) in W_k. Since all languages are countable, each W_k is countable. Now let

$$W = \cup\{W_k \mid k < \omega\}.$$

If $(T, F) \in W$ is L-saturated in $(\mathcal{L}_1, \mathcal{L}_2)$, define $D_{(T,F)} = D(\mathcal{L}_1) = D(\mathcal{L}_2)$, for any (T', F') in W let

$(T, F)R(T', F')$ iff $(D_{(T,F)} \subseteq D_{(T',F')}$ and $\forall A \in \mathcal{L}_1(\Box A \in T \Rightarrow A \in T')$ and $\forall B \in \mathcal{L}_2(\Diamond B \in F \Rightarrow B \in F'))$.

For each n-ary predicate symbol P in $\mathcal{L}_1 \cup \mathcal{L}_2$ and $c_1, \ldots, c_n \in D_{(T,F)}$ let

$$(T, F) \vDash P(c_1, \ldots, c_n) \text{ iff } (P(c_1, \ldots, c_n) \in T \text{ or } \neg P(c_1, \ldots, c_n) \in F).$$

Then $\mathbf{M} = \langle W, R, \{D_w\}_{w \in W}, \vDash \rangle$ is a Kripke model. ■

Let us prove

LEMMA 12. *Let $(T, F) \in W$ be L-saturated in $(\mathcal{L}', \mathcal{L}'')$. Then*

(i) $(\forall A \in \mathcal{L}')(A \in T \Rightarrow (T, F) \models A)$,

(ii) $(\forall B \in \mathcal{L}'')(B \in F \Rightarrow (T, F) \not\models B)$.

Proof. (i) By induction on A. Remember that A is reduced. For an atomic A, the statement holds by definition. Let A be a negated atomic sentence $\neg P(c_1, \ldots c_n)$. Then $A \in T$ implies that $P(c_1, \ldots, c_n) \notin T$, otherwise both $T \to_L \perp$ and $\perp \to_L F$ would be true, so (T, F) would be L-separable. Moreover, $A \notin F$ since (T, F) is saturated. So $(T, F) \models \neg P(c_1, \ldots, c_n)$.

Let $A = A_1 \& A_2 \in T$. Then $T \to_L A_1$ and $T \to_L A_2$, so $A_1 \in T$ and $A_2 \in T$. By the induction hypothesis $(T, F) \vDash A_1 \& A_2$.

Let $A = A_1 \vee A_2 \in T$. Then $A_1 \in T$ or $A_2 \in T$ since (T, F) is saturated. By the induction hypothesis, $(T, F) \vDash A_1 \vee A_2$. The case $A = \exists x A_2$ can be considered by analogy.

Let $A = \forall x A_1(x)$. Then $A_1(c) \in T$ for each $c \in D(\mathcal{L}')$ since $T \to_L A_1(c)$ and $A_1(c) \in \mathcal{L}'$. By induction hypothesis $(T, F) \vDash A_1(c)$ for each $c \in D((T, F))$, so $(T, F) \vDash \forall x A_1(x)$.

Let $A = \Box A_1 \in T$. The $(T, F)R(T', F')$ implies $A_1 \in T$ and $(T', F') \vDash A_1$. So $(T, F) \vDash \Box A_1$.

At last, let $A = \Diamond A_1 \in T$. By construction of the model, we have $(T, F) \in W_k$ for some k. Then there exists a pari $(T', F') \in W_{k+1} \subseteq W$ such that $(T, F)R(T', F')$ and $A_1 \in T'$. By the induction hypothesis, $(T', F') \vDash A_1$ so $(T, F) \vDash \Diamond A_1$. It proves (i) of the lemma.

(ii) Can be proved by analogy. ■

It follows from Lemma 12 that $w = (T_0, F_0)$ satisfies the condition: $w \models A$ for any $A \in \Gamma$ and $w \models \neg B$ for any $B \in \Delta$. It proves the theorem in case $L =$KQ. To complete the proof, we need

LEMMA 13. $\mathbf{W} = \langle W, R, \{D_w\}_{w \in W} \rangle$ *is an L-frame.*

Proof. If L contains an axiom $\Diamond\top$, then for any pair (T, F) in W we have $\Diamond\top \in T$, so there exists a (T', F') in W such that $(T, F)R(T', F')$; therefore, $\mathbf{M}$ is a D-model.

Let L contain an axiom $(\Box p \to p)$. Then R is reflexive. Indeed, $\Box A \to_L A$ and $B \to_L \Diamond B$ for any formulas A and B. Therefore, for each (T, F) in W we have $\Box A \in T \Rightarrow A \in T$ and $\Diamond B \in F \Rightarrow B \in F$, and then $(T, F)R(T, F)$.

Let L contain an axiom $(\Box p \to \Box\Box p)$. Prove that R is transitive. Take any $(T, F), (T', F'), (T'', F'')$ in W such that $(T, F)R(T', F')$ and $(T', F')R(T'', F'')$. Then

$$
\begin{aligned}
&D((T, F)) \subseteq D((T', F')) \subseteq D((T'', F'')),\\
&\Box A \in T \Rightarrow \Box\Box A \in T \Rightarrow \Box A \in T' \Rightarrow A \in T'',\\
&\Diamond B \in F \Rightarrow \Diamond\Diamond B \in F \Rightarrow \Diamond B \in F' \Rightarrow B \in F''
\end{aligned}
$$

So $(T, F)R(T'', F'')$. It completes the proof of Lemma 13 and of Theorem 11. ■

Now we can prove the Lyndon interpolation property for our logics.

THEOREM 14 (Interpolation). *Let L be any of quantified modal logics* KQ, DQ, TQ, K4Q, D4Q *and* S4Q*, and let $(A \to B)$ be in L. Then there exists a formula C such that both $(A \to C)$ and $(C \to B)$ are in L and also $\Omega^+(C) \subseteq \Omega^+(A) \cap \Omega^+(B), \Omega^-(C) \subseteq \Omega^-(A) \cap \Omega^-(B)$ and $D(C) \subseteq D(A) \cap D(B)$.*

Proof. By Lemma 8, one can assume that A and B are reduced. Let us replace any individual constant, which occurs A and does not occur in B, by a new variable. Then bound all free variables of the premise, which have no occurrences in the conclusion, by existence quantifiers, so we get a formula A' such that $(A \to A')$ is in L. By duality, we obtain a formula B' replacing the constants of B, which are not contained in A, by new variables and bounding them by universal quantifiers. So we have $L \vdash (B' \to B), D(A') = D(B')$ and $L \vdash (A' \to B')$.

Denote by $\mathcal{L}'$ the language satisfying $D(\mathcal{L}') = D(A'), \Omega^+(\mathcal{L}') \subseteq \Omega^+(A') = \Omega^+(A)$, $\Omega^-(\mathcal{L}') \subseteq \Omega^-(A') = \Omega^-(A)$ and by $\mathcal{L}''$ the language satisfying $D(\mathcal{L}'') = D(B') = D(A'), \Omega^+(\mathcal{L}'') \subseteq \Omega^+(B') = \Omega^+(B), \Omega^-(\mathcal{L}'') \subseteq \Omega^-(B') = \Omega^-(B)$.

Assume that there is no formula C such that both $(A' \to C)$ and $(C \to B')$ are in L and also $C \in \mathcal{L} = \mathcal{L}' \cap \mathcal{L}''$. Then the pair $(\{A'\}, \{B'\})$ is L-inseparable in $(\mathcal{L}', \mathcal{L}'')$. By Theorem 2.7 there exists an L-model $\mathbf{M} = (W, R, \{D_w\}_{w \in W}\rangle, \models)$ and an element $w \in W$ such that $w \models A'$ and $w \models \neg B'$. But it contradicts to $L \vdash (A' \to B')$.

So there is an interpolant C of $(A' \to B')$ in L. Then C is also an interpolant of $(A \to B)$. ∎

Lyndon's interpolation, and also Craig's interpolation and the Beth property do not hold in quantified modal logics containing the Barcan formula [Fine, 1979], in particular, it fails in the quantified S 5. On the contrast, the propositional logic $S5$ has the Lyndon interpolation property. It follows from Theorem 2.10 that the propositional logics K, D, T, K4, D4 and S4 have Lyndon interpolation property.

We note that Theorem 11 implies the strong completeness theorem for all systems considered.

THEOREM 15. *Let L be any of quantified modal logics* KQ, DQ, TQ, K4Q, D4Q *and* S4Q.

(i) *Let T be a set of sentences and A a sentence such that for any model* $\mathbf{M}$ *based on an L-frame $\langle W, R, \{D_w\}_{w \in W}\rangle$ and for any $w \in W, w \vDash t$ implies $w \vDash A$. Then $T \to_L A$.*

(ii) *Every L-valid formula is provable in L.*

Proof. (i) Assume that $T \to_L A$ does not hold. Then the set $T \cup \{\neg A\}$ is L-consistent, and so is the set $T' \cup \{(\neg A)'\}$ consisting of reduced formulas equivalent to the formulas of $T \cup \{\neg A\}$. It follows that the pair $(T' \cup \{(\neg A)'\}, \{\bot\})$ is L-inseparable. By Theorem 2.7, it is satisfiable in some L-frame, which contradicts the condition.

(ii) Let a formula A be L-valid. Then $\forall x_1, \ldots, \forall x_n A$ is L-valid, where $x_1, \ldots, x_n$ are all free variables of A. It follows from (i) that $\forall x_1 \ldots \forall x_n A$ is provable in L. Hence, A is provable in L. ∎

One can easily see that a propositional logic has LIP, CIP or B1 whenever its quantified extension has the same property. Indeed, we can consider propositional variables as 0-place predicates. Then an interpolant (or an explicit definition), which is a first-order formula in a langauge with 0-place predicates and without equality, is equivalent to a quantifier-free formula.

In particular, the propositional logics K, D, T, K4, D4 and S4 have Lyndon interpolation property. Although the interpolation fails for quantified S5, LIP holds for propositional S5.

2.3 Overview of interpolation and definability in modal and intuitionistic logics

This section is mainly devoted to modal logics and to extensions of the intuitionistic logic. Such extensions are called *superintuitionistic* logics. Superintuitionistic logics which are contained in the classical logic are said to be *intermediate*. An intermediate propositional logic is the same as a consistent superintuitionistic logic; it is not true for predicate logics.

We give a brief survey of basic results presented in [Gabbay and Maksimova, 2005].

The logics are presented here in different ways. First a well known Hilbert-style axiomatization of the most known propositional modal systems K, D, T, K4, D4, S4, S5 and also of the intuitionistic logic is given. Their predicate extensions are also investigated. For each of the calculi its relational semantics originated from Hintikka and Kripke is considered. For example, K is characterized by all Kripke models, K4 by models with transitive admissibility relation and S4 by reflexive and transitive models. For all the mentioned basic logics the strong completeness theorems guarantee the equivalence of the syntactic consequence relations to the corresponding semantical consequence relations defined by Kripke models.

On the other hand, it is well known that there are modal and intermediate logics which are Kripke-incomplete. When we are interested in general study of large families of logics, for instance, the family of all superintuitionistic logics or all normal extensions of the minimal normal modal logic K, algebraic methods are very fruitful. With every normal modal logic a variety of modal algebras is associated, and also there is a one-to-one correspondence between superintuitionistic logics and varieties of Heyting algebras. Strong algebraic completeness of all propositional modal and superintuitionistic logics makes possible to investigate these logics in an uniform way. The algebraic semantics is closely connected with relational semantics via representation theorems.

It is well known that the intuitionistic logic Int may be translated into S4. This translation was introduced by Gödel in order to give an interpretation of Int in the classical logic with an additional provability operator. Provability interpretation of modalities stimulated interest to the provability logic G. Not long ago an interpretation of S4 as a logic of proofs was found [Artemov, 2001]. Also the logics S4 and Int have a natural topological interpretation.

2.3.1 Intermediate logics and extensions of S4

In Chapter 3 of our book [Gabbay and Maksimova, 2005] we pay special attention to extensions of the intuitionistic logic and of the modal logic S4. An algebraic semantics and also Kripke semantics for these logics is presented in more detail. We explain the main interrelations of the family $E(\text{Int})$ of superintuitionistic logics and of the family $NE(\text{S4})$ induced by Gödel's translation. In particular, any intermediate logic L has an infinite family of its modal companions in $NE(\text{S4})$, which have L as their superintuitionistic fragment.

2.3.2 Interpolation and joint consistency in the classical and intuitionistic logic

The study of interpolation begins with Chapter 4. Interpolation theorem for the classical predicate logic was discovered by W. Craig in 1957. It says that if A implies B, where each of A and B has its own language, then there is an interpolant, i.e., a formula C in the common language such that A implies C and C implies B. At the same time A.Robinson proved his joint consistency theorem which appeared to be equivalent to Craig's interpolation theorem. Robinson's theorem states that the join of two first order theories is consistent if their intersection is a complete theory in the common language. Lyndon's interpolation theorem proved in 1959 is a strengthening of Craig's theorem and takes into consideration also negative and positive occurrences of shared predicate symbols. Beth's theorem on implicit definability proved in 1953 says that any predicate implicitly definable in a first order theory is explicitly definable.

Some analog's of these statements are valid also for the intuitionistic and the most known modal logics.

Craig's interpolation is equivalent to Robinson's joint consistency, and Lyndon's interpolation theorem is proved for the classical predicate logic. The general form of Robinson's consistency property RCP fails in the intuitionistic predicate logics HQ; a weaker form of RCP is equivalent to Craig's interpolation property CIP and holds in HQ, a semantic proof originated from [Gabbay, 1971] is given. On the other hand, we prove that in propositional intermediate logics the general form of RCP is equivalent to CIP. A derivation of Beth's property from CIP is given for the intuitionistic predicate logic. Also Kreisel's proof of validity of the Beth property for any propositional intermediate logic is presented. Note that there are intermediate predicate logics without Beth's property.

2.3.3 Lyndon's interpolation in modal logics

In Section 2.2 we already presented a proof of Lyndon's interpolation property LIP for quantified extensions of basic modal logics $\mathrm{K}, \mathrm{T}, \mathrm{D}, \mathrm{K4}$ and S4 and for some others. Also the propositional S5 has LIP. At the same time, the quantified extension of S5, as well as other systems satisfying the Barcan formula, have neither Lyndon's nor Craig's interpolation, nor Beth's property; K.Fine's proof is presented in [Gabbay and Maksimova, 2005, Chapter 5].

Some examples of propositional modal logics which have CIP but do not possess LIP are found. A proof of Craig's interpolation property is given for a number of propositional modal logics including Grzegorczyk's logic Grz, its extension Grz.2 and the provability logic G.

Section 5.5 of [Gabbay and Maksimova, 2005] deals with propositional logics. We define a class of so-called L-conservative formulas, which can be added to L as new axiom schemes without loss of interpolation. Also we prove that the interpolation properties are preserved when we go from predicate logics without equality to their extensions with equality.

2.3.4 Interpolation in intermediate logics

Chapters 6–12 of [Gabbay and Maksimova, 2005] are devoted to interpolation and definability problems in propositional logics. Of course, the results have immediate applications in predicate logics: if the propositional fragment of a predicate logic L lacks interpolation or Beth's property, so does L itself. The main results of these chapters concern intermediate logics and also modal logics over S4, G or K4.

Chapter 6 of our book contains full description of superintuitionistic logics with Craig's interpolation property obtained by [Maksimova, 1977]. It turns out that in the continuum of intermediate logics, only seven have have Craig's interpolation. All of them are finitely axiomatizable and have the finite model property. For the proof, the algebraic semantics via varieties of Heyting algebras is used and the equivalence of CIP in a logic L to amalgamability of the corresponding variety $V(L)$ is stated (a class of algebras is amalgamable if any two algebras with a common subalgebra have their common extension). We also prove that the interpolation problem over the intuitionistic logic Int is decidable: for any finite set Ax of axiom schemes to determine, whether the calculus Int $+$ Ax has CIP; also the amalgamation problem is base-decidable for varieties of Heyting algebras.

2.3.5 Interpolation and definability in modal logics

It is necessary to note that the definitions of interpolation and of Beth's property essentially depend on the consequence relation in the logic under consideration. When we go to modal logics, we have at least two natural logical consequence relations: provable implication and deducibility. They are equivalent in superintuitionistic logics due to deduction theorem but not equivalent in normal modal logics, where only a weaker form of the deduction theorem holds. So in modal logics we consider two forms of interpolation: CIP and IPD (interpolation property for deducibility), and two forms B1 and B2 of the Beth property. It is shown in [Gabbay and Maksimova, 2005, Chapter 7] that B1 is equivalent to CIP for all modal logics but all other properties are not equivalent. CIP implies IPD, B1 implies B2, and IPD and B2 are independent. A full diagram of interrelations of these properties as well as their algebraic equivalents are found. In particular, IPD is equivalent to the amalgamation property AP, CIP to the superamalgamation property SAP and B2 to epimorphisms surjectivity ES* of the corresponding variety of modal algebras.

2.3.6 Interpolation over S4

It appears that the behavior of interpolation over S4 is similar to interpolation in superintuitionistic logics. It is shown in [Gabbay and Maksimova, 2005, Chapter 8] that all modal logics with IPD in NE(S4) are modal companions of superintuitionistic logics with CIP but there is an intermediate logic with CIP that has no modal companions with IPD. On the other hand, all modal companions of intermediate logics with CIP have a weaker version of interpolation, which is CIP restricted to those formulas $A \to B$, where all occurrences of variables are preceded by necessity symbol.

It is proved by Maksimova [Maksimova, 1979], that there are only finitely many modal logics with IPD in the family NE(S4). The list of 49 logics is found, which contains all extensions of S4 with IPD, in this list there are 12 logics that have IPD but do not have CIP, and CIP is proved for 31 logics. All the 49 logics are finitely axiomatizable and have the finite model property.

The problem of interpolation is completely solved for extensions of the Grzegorczyk logic Grz and for those logics over S4 which are well-representable by Kripke frames, in particular, for logics over S5. We leave open the problem how many logics over S4 have CIP or IPD. In fact, interpolation problem is still open for 6 normal extensions of S4.

Nevertheless, it is proved that IPD and CIP problems are decidable over S4, and amalgamation and superamalgamation are base-decidable in varieties of closure algebras. Complexity bounds for interpolation, amalgamation and some other problems over Int and S4 are found. For example, the interpolation problems over Int and Grz are PSPACE-complete.

2.3.7 Interpolation and the Beth property over K4

In the next two chapters of [Gabbay and Maksimova, 2005] the family NE(K4) of normal modal logics containing K4 is investigated. We find a strong necessary condition for interpolation which implies failure of interpolation for a large family of logics over K4. In particular, all infinite-slice extensions of K4.3 do not possess interpolation.

On the contrast, a strong positive result holds for extensions of K4. We prove that all logics in NE(K4) have the Beth property B2. An algorithm for constructing explicit definitions is found for logics of finite slices characterized by transitive and antisymmetric Kripke models.

2.3.8 Extensions of the provability logic

It is evident that all the results concerning logics over K4 are applicable to all logics over S4 or over the provability logic G since all these logics contain K4. The family NE(G) of normal extensions of G is studied in [Gabbay and Maksimova, 2005, Chapter 12]. Here we see that the results on interpolation over S4 can not be extended to all modal logics. The picture of interpolation over G is quite different from that of NE(S4), where only finitely many logic possessed interpolation and all of them were finitely axiomatizable and had the finite model property.

We build a continuum of normal modal logics with CIP. We find logics with CIP which are neither finitely axiomatizable nor finitely approximable. The most interesting of these logics is a logic Gγ constructed in [Gabbay and Maksimova, 2005, Section 12.1]. It appeares that Gγ is the greatest among the infinite-slice logics with IPD in NE(G); in addition, it is decidable. Also we prove that IPD is equivalent to CIP in all finite-slice logics over G, it was not true in NE(S4).

2.3.9 Syntactic proofs of interpolation

In Chapters 3 and 6–12 of [Gabbay and Maksimova, 2005] algebraic methods are developed, that allows to formulate and prove our results in two areas: in logic and in algebra. Actually even in "algebraic" chapters we apply relational semantics in parallel with algebraic one, where it is possible. In particular, we propose semantical methods for proving interpolation. The most of the results of Chapters 6–12 of our book are formulated also in the language of Kripke semantics.

In Chapter 4 (as well as in Chapter 5) of [Gabbay and Maksimova, 2005] semantic methods are used for proving interpolation. On the one hand, we prove an extension of Model Existence Theorem which implies Strong Completeness Theorem as a partial case. But on the other hand, semantic methods do not give any algorithm for constructing an interpolant. In syntactic proofs special Gentzen-style or tableaux calculi are required which have convenient rules of inference and admit cut elimination. Then an interpolant is constructed from a derivation of the formula $A \to B$ or of the sequent $A \to B$. In [Gabbay and Maksimova, 2005, Chapter 13] we give a syntactic proof of Lyndon's interpolation for the intuitionistic (and also for the classical) predicate logic by modifying Schütte's proof [Schütte, 1962] of interpolation for HQ.

The proof also works for the most of fragments of HQ. For the fragment containing neither disjunction nor existence quantifier, a weaker form of Lyndon's interpolation is proved. Note that this fragment coincides with the $\{\rightarrow, \perp, \&, \forall\}$-fragment of the intuitionistic logic of finite domains and also with the analogous fragment of the logic of constant domains.

Another proof of weak Craig's interpolation for $\{\rightarrow, \perp, \&, \forall\}$-fragment of HQ is presented in [Gabbay and Maksimova, 2005, Chapter 15], which deals with interpolation in intuitionistic logic programming (see [Gabbay and Olivetti, 2005]). A counter-example to the general form of Craig's interpolation is given in [Gabbay and Maksimova, 2005, Section 15.5].

The aim of Chapter 16 of [Gabbay and Maksimova, 2005] on goal-directed proof systems (see also [Gabbay and Olivetti, 2000]) is to study interpolation properties for implicational fragment of a variety of substructural, strict modal and intuitionistic and intermediate logics. The methodology is proof-theoretical and makes use of a goal directed formulation of these fragments which follows the logic programming style of deduction. We obtain more refined as well as new kinds of interpolation theorems for our logics and investigate new global methods for obtaining interpolation.

2.3.10 Interpolation by translation

Interpolation by translation first proposed in [Gabbay and Ohlbach, 1992] gives some uniform algorithmic methodology for finding interpolants. It operates with translations of non-classical logics into classical first order teories and introduces so-called expansion interpolation. That allows to find interpolants in the classical theories using the existing algorithms and then translate them back into non-classical theories. In [Gabbay and Maksimova, 2005] two examples from modal logic are considered: quantified S5 and propositional S4.3. These logics lack ordinary interpolation but have expansion interpolation.

It is impossible to give the details and all the references in this short chapter. The reader may find them in our book [Gabbay and Maksimova, 2005]. In the next section we discuss further directions of research and some results which were not included in this volume.

3 FURTHER RESULTS AND DISCUSSION

This section contains two subsections. One describing additional material already existing in papers, and one describing the new research / challenges that need to be addressed.

3.1 Further results

In [Gabbay and Maksimova, 2005] the picture of methods and results was presented which gives some basic knowledge on interpolation and definability in most popular logics. At the moment we have a lot of material which was not included, and we are planning to write volume 2 of this book. In this section we continue our exposition of further directions and known results on interpolation and definability in non-classical logics.

The algebraic approach to modal logic can be applied to other logics. For instance, the same algebraic criteria for interpolation and definability properties are valid for multi-modal and (sometimes with some changes) for various non-classical logics. It allows to prove or disprove interpolation and Beth's properties in large families of logics.

The significance of our negative results (criteria for failure of interpolation or Beth properties) is twofold: 1) if we wish to find a logic for solving some tasks, we prefer to choose a logic having the mentioned good properties; 2) there is a general method for constructing counter-examples to interpolation in particular logics, this may be helpful for understanding which logical constants should be added in order to obtain the desirable properties. Actually this method is only performed in volume 1, and must be developed in more detail.

3.1.1 Temporal logics

First we consider temporal logics.

Temporal logics form a natural class of multi-modal logics. In Maksimova [Maksimova, 1991b], [Maksimova, 1991a], we proved failure of CIP, IPD and of the Beth properties in temporal logics with discrete (linear or branching) time. A counter-examples are found. The same counter-examples work for CTL and related logics.

F. Wolter [Wolter, 1997] proved that tense logics of linear time have no interpolation. On the other hand, the modal μ-calculus has uniform interpolation [D'Agostino and Hollenberg, 1998].

3.1.2 Beth properties and epimorphisms surjectivity

3.1.2.1 Projective Beth properties. We consider various versions of the Beth definability property for propositional nornal modal logics, and also for superintuitionistic and relevant logics. We discuss interrelations of these properties, and find their algebraic equivalents in case of modal and superintuitionistic logics.

The Beth properties B1 and B2 are particular cases of projective Beth's properties PB1 and PB2 defined as follows.

Let $\mathbf{x}$, $\mathbf{q}$, $\mathbf{q}'$ be disjoint lists of variables not containing y and z, $A(\mathbf{x}, \mathbf{q}, y)$ a formula. We say that a logic L has *the projective Beth property* PB1 if

$\vdash_L A(\mathbf{x}, \mathbf{q}, y) \& A(\mathbf{x}, \mathbf{q}', z) \to (y \equiv z)$ implies $\vdash_L A(\mathbf{x}, \mathbf{q}, y) \to (y \equiv B(\mathbf{x}))$ for some formula $B(\mathbf{x})$.

Further, L has the projective Beth property PB2 if

$A(\mathbf{x}, \mathbf{q}, y), A(\mathbf{x}, \mathbf{q}', z) \vdash_L y \equiv z$ implies $A(\mathbf{x}, \mathbf{q}, y) \vdash_L y \equiv B(\mathbf{x})$ for some $B(\mathbf{x})$.

In [Gabbay and Maksimova, 2005, Chapter 7] a diagram of interrelations between the properties B1, B2 and interpolation properties CIP and IPD was found for normal modal logics. It was proved that B1 is equivalent to CIP, and implies B2 and IPD; the properties IPD and B2 are independent.

The following relations are stated in Maksimova [Maksimova, 1999b]:

$$\text{PB1} \iff \text{B1} \iff \text{CIP} \Rightarrow \text{B2+IPD} \Rightarrow \text{PB2} \Rightarrow \text{B2}.$$

Since all normal extensions of K4 have B2 (see [Gabbay and Maksimova, 2005, Chapter 11]), IPD implies PB2 in logics over K4.

3.1.2.2 Algebraic equivalent of PB2. It is proved in Maksimova [Maksimova, 1999b] that a modal logic has PB2 if and only if its corresponding variety has strong epimorphisms surjectivity SES:

For any $\mathbf{A}$,$\mathbf{B}$ in V(L), for any monomorphism $\alpha : \mathbf{A} \to \mathbf{B}$ and for any $x \in \mathbf{B} - \alpha(\mathbf{A})$ there exist $\mathbf{C} \in V$ and monomorphisms $\beta : \mathbf{B} \to \mathbf{C}$ and $\gamma : \mathbf{B} \to \mathbf{C}$ such that $\beta\alpha = \gamma\alpha$ and $\beta(x) \neq \gamma(x)$.

It was shown in [Gabbay and Maksimova, 2005, Chapter 7] that a modal logic has B2 if and only if its corresponding variety has epimorphisms surjectivity ES*. Another form BP of the Beth property equivalent to a general form of epimorphisms surjectivity was considered in Sain [Sain, 1989]. In our terms it can be formulated as follows: Let P, Q and Q' be pairwise disjoint sets of variables and $Q' = \{q_i' | \ q_i \in Q\}$. If $\Gamma(P, Q), \Gamma(P, Q') \vdash_L (q_i \equiv q_i')$, then for every $q_i \in Q$ there exists a formula $B_i(P)$ such that $\Gamma(P, Q) \vdash (q_i \equiv B_i(P))$.

It is easy to see that PB2 $\Rightarrow$ BP $\Rightarrow$ B2 for any variety. If the set Q is finite, BP is equivalent to B2 [Maksimova, 1999b], [Hoogland, 2000]. It is not clear if the general form of BP is equivalent to B2 or PB2 in the case of modal logics.

Convenient algebraic criteria for validity of PB2 and SES are found in Maksimova [Maksimova, 1999b], [Maksimova, 2003b].

3.1.3 Projective Beth property over Int

In intermediate logics two versions PB1 and PB2 of projective Beth property are equivalent due to deduction theorem, so we write PBP instead of PB2. It is easy to derive PBP from CIP by analogy with the proof of B1 from CIP. The equivalence of PBP to SES also holds for intermediate logics [Maksimova, 1999a].

It appeared that the behavior of PBP over Int is similar to that of interpolation. In Maksimova [Maksimova, 2000b] we have obtained full description of propositional superintuitionistic logics with the projective Beth property. There are exactly 16 logics with PBP over Int. All of them are finitely axiomatizable and have the finite model property. Their Kripke characterization is rather simple.

Decidability of projective Beth property over Int is proved in [Maksimova, 2001]. It follows that strong epimorphisms surjectivity is base-decidable in varieties of Heyting algebras.

The methods and results concerning intermediate logics have immediate applications in positive and paraconsistent logics and also in modal logics.

3.1.4 Positive and paraconsistent logics

3.1.4.1 Positive logics. In the language of the intuitionistic logic there is a formula $\perp$ which implies any formula in Int, and negation is expressed by $A \to \perp$. If $\perp$ and $\neg$ are ejected, we come to positive logics. Let us consider positive logics containing positive fragment Int^+ of the intuitionistic logic. Each positive logic can be considered as positive fragment of a suitable superintuitionistic logic. But in general interpolation and projective Beth's property are not preserved by transfer to positive fragments. We define a special translation of logics over Int^+ into intermediate logics, which allows to apply the results of the previous section to positive logics.

Full description of positive logics with interpolation or projective Beth property is found. There are exactly 7 positive logics with PBP, among them 4 logics have CIP and the others do not have.

Moreover, we prove that interpolation and projective Beth's property are decidable over Int^+. The problem of interpolation is PSPACE-complete over Int^+, and so is PBP problem. One can find proofs in [Maksimova, 2002c], [Maksimova, 2003a].

3.1.4.2 Applications to paraconsistent extensions of Johansson's minimal logic The methods of studying positive and intermediate logics also applicable to extensions of the minimal logic J which has the same positive fragment as the intuitionistic logic but is a good base for paraconsistency. The constant $\perp$ has no special features. In particular, the formula $\perp \to p$ is not a theorem of J. Many results can be transferred from positive logics to extensions of J although still we have no solution for the problems of interpolation or projective Beth's property over J. It is open problem how many logics over J have interpolation or PBP.

The papers [Maksimova, 2002a], [Maksimova, 2003a] present current knowledge on the subject.

3.1.4.3 Some remarks on relevant logics It was proved by Urquhart [Urquhart, 1993] that the basic relevant logics, among them E and R, have neither CIP nor Beth definability properties. Nevertheless, CIP holds in OR which is R without distributivity axiom [McRobbie, 1983].

Only some weak forms of the deduction theorem hold in relevant logics. On this reason, the interrelations of different forms of interpolation and Beth property given in Section 3.1.2 (and also in Section 3.1.6 below) are broken. The usual implication from CIP to IPD holds for extensions of the logic E of entailment if the language includes a propositional constant **t** ("the strongest truth"), where $\vdash_L$ denotes the deducibility with modus ponens and adjunction. To preserve a standard proof of the Beth property from CIP, we need an intensional conjunction $\circ$ (that is commutative and associative) and the following definition

PB1'. If $\vdash_L A(\mathbf{x}, \mathbf{q}, y) \circ A(\mathbf{x}, \mathbf{q}', z) \to (y \leftrightarrow z)$ then $\vdash_L A(\mathbf{x}, \mathbf{q}, y) \to (y \leftrightarrow B(\mathbf{x}))$ for some $B(\mathbf{x})$.

Thus CIP implies PB1' in extensions of the relevance logic R. The equivalence of IPD to the amalgamation property holds for all extensions of E [Czelakowski, 1982], and the equivalence of CIP in L to the superamalgamation in its associated variety V(L) holds for extensions of R (or of the fragment of R with **t**, &, $\to$ and $\circ$).

3.1.5 Modal logics and projective Beth property

In modal logics as well as in intermediate logics there is a similarity of projective Beth property PB2 to interpolation properties.

1. It was proved in [Gabbay and Maksimova, 2005, Chapter 10] that for any infinite slice logic with IPD over K4, its reflexive fragment is contained in Grz.2. The theorem remains true if IPD is replaced with PB2 (which is implied by IPD over K4) [Maksimova, 2002b].

2. NE(S4) versus E(Int). The results of Section 3.1.3 are applied in study of modal logics over S4 [Maksimova, 2004]. We find some similarity and some difference of modal and intermediate logics with respect to PB2. In particular, any modal logic over S4 with PB2 is a modal companion of an intermediate logic with PB2. On the other hand, there are three intermediate logics with PB2 which have no modal companions with PB2.

We proved that there are exactly 16 superintuitionistic logics with PB2. At this moment we have no list of logics with PB2 in NE(S4). And we do not know if their number is finite.

3. Full description of logics with PB2 over Grz and over S5 is found. There are exactly 13 normal extensions with PB2 over Grz and 4 over S5. The property PB2 is decidable over Grz and over S5. Some complexity bounds are found.

3.1.6 Restricted interpolation and restricted amalgamation

Interconnections between various versions of Craig's and Beth's properties essentially depend on the form of deduction theorem which holds (or does'not hold) in the logic under consideration. The same is true for categorical properties such as congruence extension property and various forms of amalgamation and epimorphisms surjectivity in the corresponding class of algebras. It may happen (for instance, in some substructural logics) that a logic has CIP but does not have IPD.

Some natural forms of interpolation and of Robinson property are relatively new and will be a subject of research. In [Maksimova, 2003b] we introduced a *restricted interpolation property*:

IPR. If $A(\mathbf{p},\mathbf{q}), B(\mathbf{p},\mathbf{r}) \vdash_L C(\mathbf{p})$, then there exists a formula $A'(\mathbf{p})$ such that $A(\mathbf{p},\mathbf{q}) \vdash_L A'(\mathbf{p})$ and $A'(\mathbf{p}), B(\mathbf{p},\mathbf{r}) \vdash_L C(\mathbf{p})$.

We proved that the restricted version IPR of interpolation property is equivalent to the restricted amalgamation in varieties of modal or Heyting algebras and showed that IPR follows from PB2. In addition we found an algebraic criterion for PB2 with use of the restricted amalgamation property. A class V has *Restricted Amalgamation Property* if it satisfies the condition:

RAP. For each $\mathbf{A},\mathbf{B},\mathbf{C} \in V$ such that $\mathbf{A}$ is a common subalgebra of $\mathbf{B}$ and $\mathbf{C}$ there exist an algebra $\mathbf{D}$ in V and homomorphisms $\delta : \mathbf{B} \to \mathbf{D}$, $\varepsilon : \mathbf{C} \to \mathbf{D}$ such that $\delta(x) = \varepsilon(x)$ for all $x \in \mathbf{A}$ and the restriction δ' of δ onto $\mathbf{A}$ is a monomorphism.

THEOREM 16 (Maksimova, 2003b). *Let L be a normal modal logic. Then L has* PB2 *if and only if $SI(V)$ has* RAP *and $FI(V)$ has* SES*.*

We note that V has RAP iff $SI(V)$ has RAP, so PB2 implies IPR.

Interconnections between syntactical and categorical properties of equational theories are established in [Maksimova, 2003c]; in [Maksimova, 2003b] modal logics are discussed. We obtain the following relations for all normal modal logics and varieties of modal algebras:

(1) IPD + B2 $\iff$ StrAP,
(2) IPD + B2 $\Rightarrow$ PB2 $\iff$ SES,
(3) PB2 $\not\Rightarrow$ IPD $\iff$ AP, IPD $\not\Rightarrow$ PB2, PB2 $\Rightarrow$ B2,
(4) PB2 $\Rightarrow$ IPR $\iff$ RAP,
(5) IPD $\Rightarrow$ IPR, IPR + B2 $\not\Rightarrow$ IPD,
(6) IPR $\not\Rightarrow$ B2, B2 $\not\Rightarrow$ IPR.

We do not know if IPR+B2 implies PB2 in modal logics. We can prove that it holds for positive logics of Subsection 3.1.4.1.

In fact, most of the relations (1)–(6) are valid for arbitrary varieties which are congruence-distributive, have congruence extension property and satisfy an additional condition: any subalgebra of a subdirectly irreducible algebra is finitely indecomposable. The proofs are given in [Maksimova, 2003b]. In particular, all of (1)–(6) hold for substructural logics without contraction but with weakening and exchange; they are characterized by suitable varieties of residuated lattices investigated in [Ono, 2003].

All intermediate and positive logics, and also modal logics over K4 have the Beth property B2. For those logics and corresponding varieties we obtain:

IPD $\iff$ AP $\iff$ StrAP,
IPD $\Rightarrow$ PB2 $\iff$ SES,
PB2 $\not\Rightarrow$ IPD,
PB2 $\Rightarrow$ IPR $\iff$ RAP.

Note that IPR implies CIP in extensions of S5, therefore, all properties PB1, PB2, B1, CIP, IPD, IPR are equivalent over S5. The results of this section are published in Maksimova [Maksimova, 2003b], [Maksimova, 2003c].

Problem. *The implication from* PB2 *to* IPR *in modal logics is proved by purely algebraic argument. It would be interesting to find a direct proof in logical terms.*

3.1.7 Variable separation

3.1.7.1 Principles of variable separation in relevant and substructural logics. Relevant calculi satisfy so-called Relevance principle: If the implication $A \to B$ is valid then A and B share a variable. It was proved that the the most known relevant systems such as the logic R of relevant implication and the logic E of entailment are undecidable [Urquhart, 1984] and have neither interpolation nor the Beth property [Urquhart, 1993], although some relevant logics have interpolation [McRobbie, 1983].

In Maksimova [Maksimova, 1976], [Maksimova, 1982b] we formulated and proved various principles of variable separation, which are in some sense similar to interpolation, for the most known relevant systems R of relevant implication and E of entailment. For example, the following statement holds:

PVS if two formulas $(A_1 \to A_2)$ and $(B_1 \to B_2)$ have no variable in common and the formula $((A_1 \& B_1) \to (A_2 \vee B_2))$ is provable in E (or R), then at least one of $(A_1 \to A_2)$ and $(B_1 \to B_2)$ is also provable.

It is clear that PVS implies Hallden's property

HP If $A \vee B$ is provable and A and B have no variable in common then A or B is also provable.

Naruse [Naruse *et al.*, 1998] considered PVS in substructural logics.

3.1.7.2 Hallden property and variable separation in modal and intermediate logics, and joint embedding property in varieties. It was shown (see [Gabbay and Maksimova, 2005, Section 4.3]) that CIP implies PVS in intermediate logics; also one can prove that IPR implies PVS. Connections of PVS in intermediate logics with Hallden's property of their modal companions were studied in Chagrov and Zakharyaschev [Chagrov and Zakharyaschev, 1993], [Chagrov and Zakharyaschev, 1997]; it was proved that HP and PVS are undecidable over Int and over S4.

It is clear that in modal logics, PVS and HP are equivalent; CIP implies HP in extensions of D but the provability logic G has CIP and does not possess HP.

Algebraic equivalents of these properties were found in [Maksimova, 1995], where some other related properties were considered. It appeared that HP and PVS are in close relation with joint embedding property of corresponding varieties of algebras. In particular, HP in a modal logic L is equivalent to

JSEP for any $\mathbf{A}, \mathbf{B}$ in $V(L)$ there is a $\mathbf{C} \in V(L)$ and embeddings $\varphi : \mathbf{A} \to \mathbf{C}$, $\psi : \mathbf{B} \to \mathbf{C}$ such that $\varphi(a) \not\leq \psi(b)$ for any $a \in \mathbf{A}, a \neq \bot$, $b \in \mathbf{B}, a \neq \top$.

HP and related properties in intermediate logics were studied in [Wronski, 1976] and [Suzuki, 1990].

3.1.8 Decidable properties of logics and of varieties

3.1.8.1 Decidable and strongly decidable properties of logical calculi. We already mentioned that interpolation is decidable over Int and over S4, the projective Beth property is decidable over Int and so on.

Decidability of interpolation over Int means that there is an algorithm, which for any finite list Ax of axiom schemes decides if the calculus Int + Ax has interpolation property. Here Int+ Ax denotes a calculus obtained from some standard calculus for the intuitionistic logic by adding new axiom schemes but not rules of inference. Decidability (or undecidability) of various properties of logical calculi was a subject of many papers (see [Chagrov and Zakharyaschev, 1997]).

We say that a property P is *strongly decidable over a logic L* if there is an algorithm, which for any finite list Rul of additional axiom schemes and rules of inference decides if the calculus $L + Rul$ has the property P. It is proved in [Maksimova, 2000a] that interpolation is strongly decidable over Int, S5 and Grz, consistency is strongly decidable over D. Strong decidability of some other properties is also shown. An overview of decidable and strongly decidable properties of logical calculi is given in [Maksimova, to appear].

3.1.8.2 Tabular logics and varieties. It is worth noting that all versions of interpolation and of the Beth property, and also Hallden property, are decidable on the class of tabular modal logics. A logic is tabular if it can be characterized by finitely many finite algebras or (equivalently) by finitely many Kripke frames. We can propose an uniform constructive method of checking interpolation and related properties through amalgamation and epimorphisms surjectivity of suitable finite classes of finite algebras. We will describe this method in volume 2. Some ideas can be found in [Maksimova, 1999c].

3.2 Further discussion

We have already mentioned that writing volume 1 of our book [Gabbay and Maksimova, 2005] has indentified some areas that need to be further summarised and investigated so that we have a better understanding and a more complete picture of interpolation. This section will describe the challenges we hope to address in volume 2. Since what we have here is hopeful thinking, and a discussion of what seems to us promising points of view, we named this section "Further discussion" as compared with the previous section "Further results."

3.2.1 Interpolation and artificial intelligence

We do not have much more to say here beyond what we have already said in Section 1.2. It is sufficient that we stress again that we view the challenges of Section 1.2 as mainstream research in logic and whatever we find and include in volume 2 will enhance our general understanding of Interpolation.

3.2.2 Interpolation for classical theories

We have discussed interpolation by translation in Chapter 14 of [Gabbay and Maksimova, 2005]. The research on this topic is only beginning and there is more to be done. There are two reasons for intensive continuation of this research. The first is that we have the general methodology of *Logic by Translation*, and so we need to look also at interpolation by translation as part of the program. The second reason is that when we translate from one logic into another, we need to have interpolation in the target logic, relative to the theory controlling the translation. This gives urgency to studying interpolation in general for arbitrary or for particular theories in the target logic and also for special fragments of the target logic, which is usually classical logic. To be concrete, we can study interpolation for interesting well known theories of classical logic. For example let FF be the theory of finite fields (i.e., defined semantically by the class of all finite fields).

Does it have interpolation?

How about the theory of all finite projective planes in some obvious language?

We will try and get some answers for some well known theories.

3.2.3 A semantic/categorial engine for interpolation

The construction we gave in [Gabbay and Maksimova, 2005, Chapter 4] proving interpolation for intuitionistic logic seems to be actually a general purpose construction for proving interpolation. M. Makkai [Makkai, 1995] has observed that the steps can be generalised and presented in the framework of categorial logics. In fact the proof in Section 4.2 shows something categorical, stronger than interpolation.

We should mention that a very elegant and conceptually interesting proof of interpolation for intuitionistic logic was given by A. Pitts [Pitts, 1983]. He uses locale-theoretic methods. Pitt's proof is constructive as opposed to our Section 4.2 which is model-theoretic. On the other hand, the proof we give can be generalised into a general engine for proving interpolation for a variety of systems. We postpone this study to volume 2. One by product of such an engine will be that for some logics for which interpolation fails, a better understanding can be gained by trying to see why the general engine cannot prove interpolation for the logic.

3.2.4 Interpolation in computer science

The important role of interpolation in theoretical computer science was first observed by M. Sadler [Maibaum and Sadler, 1985] and further developed by many others, see [Bergstra *et al.*, 1990], [Rodenburg and van Galbeek, 1988], [Dimitrakos and Maibaum, 1997b], [Dimitrakos and Maibaum, 1997a], [Ehrich, 1982], [Marx, 1995]. See also Chapters 4–5 of T. Dimitrakos [Dimitrakos, 1998] and see also the paper [Bicarregui *et al.*, 2001]. We find Veloso's paper [Veloso, 1993] most clear and our exposition will make use of it.

The notion of interest from computer science is that of the *Modularisation Theorem.* It has to do with composing specification. It is not necessary for us to go in this chapter into computer science examples. It suffices to say that if we have a language $\mathcal{L}_0$ with specification G_0 about the predicates of $\mathcal{L}_0$, then implementing this specification on some machine means in logical terms that we have a language $\mathcal{L}_2$ (of the machine) and axiom G_2 (describing the machine, how it works), and a translation $f : \mathcal{L}_0 \mapsto \mathcal{L}_2$, actually describing the implementation. We must have that all the properties of G_0 are soundly and faithfully implemented, namely for any φ in $\mathcal{L}_0$, if $G_0 \vdash_0 \varphi$ then $f(G_0) \vdash_2 f(\varphi)$, and vice versae.

In logic we have many such examples of implementations, we call them *Interpretations.* It would be useful to think of the "implementation" of modal predicate logic K4B (modal K4 with the Barcan formula, i.e., constant domains modal K4) in classical logic CQ, as our working example. This "implementation," together with ideas of Interpolation by translation [Gabbay and Ohlbach, 1992], [Gabbay and Maksimova, 2005], will be used by us to explain the mudularisation theorem and its connection with interpolation.

So let us go on with our discussion of specifications and their implementations. So assume we have specification $(\mathcal{L}_0, G_0)$ and we implement it on a machine $(\mathcal{L}_2, G_2)$. We may expand our specification in a conservative way to an additional language $(\mathcal{L}_1, G_1)$. Naturally we want to expand our implementation on the machine $(\mathcal{L}_2, G_2)$ into an expanded specification $(\mathcal{L}_3, G_3)$. The most natural requirement is that the additional implementation is conservative over the old implementation. We do not want to spoil what we already have!

Going back to logic, suppose we now extend the language $\mathcal{L}_0$ into the language $\mathcal{L}_1$ by adding new symbols $\mathcal{C}$ (thus $\mathcal{L}_1 = \mathcal{L}_0 + \mathcal{C}$) and new axioms / specification G_1. This is a very natural thing to do. A builder for example, may specity what he wants in a kitchen in language $\mathcal{L}_0$ and then realise he has to add the oven and microwave in there, so he gives additional specification. We need to assume that $\mathcal{L}_2$ and $\mathcal{L}_1$ are disjoint (otherwise life gets complicated!), We also assume that G_1 is a conservative extension of G_0, i.e., the additional specification does not change what was already agreed about $\mathcal{L}_0$.

We now want to expand the implementation language $\mathcal{L}_1$ and the axioms G_1 into $\mathcal{L}_3$ and axioms G_3 such that the additional specification $(\mathcal{L}_1, G_1)$ can be implemented (via a function $g : \mathcal{L}_1 \mapsto \mathcal{L}_3$) in $(\mathcal{L}_3, G_3)$. Of course g extends the original f. The modularisation theorem says that this can be done in a natural way and that G_3 is a conservative extension of G_2.

Figure 6 shows the situation; $i_{0,1}$ and $i_{2,3}$ are inclusion mappings.

$$\begin{tikzcd}
(\mathcal{L}_1, G_1) \arrow[r, dashed, "g"] & (\mathcal{L}_3, G_3) \\
(\mathcal{L}_0, G_0) \arrow[u, "i_{0,1}"'] \arrow[r, "f"] & (\mathcal{L}_2, G_2) \arrow[u, dashed, "i_{2,3}"']
\end{tikzcd}$$

Figure 6.

It turns out that the modularisation theorem is related (equivalent) to versions of interpolation, see [Veloso, 1993].

3.2.5 *Case study: Implementation of constant domains modal K4 in classical logic*

Let $\mathcal{L}_0$ be the language of modal predicate logic with one modality $\Box$ and let G_0 be the axioms for modal K4 with constant domains. Let $\mathcal{L}_1$ be the language of classical logic with binary relation xRy for the accessibility relation of the Kripke semantics and the domain relation $U(x,y)$, meaning y is in the domain of the world x. For each atomic predicate $P(x_1,\ldots,x_n)$ of the modal language let $P^*(x_0,x_1,\ldots,x_n)$ be a corresponding predicate of the classical language $\mathcal{L}_2$.

Let f be the following translation, indexed by a variable t:

$$f_t(P(x_1,\ldots,x_n)) = P^*(t,x_1,\ldots,x_n),$$
$$f_t(A\&B) = f_t(A)\&f_t(B),$$
$$f_t(\neg A) = \neg f_t(A),$$
$$f_t(\forall xA(x)) = \forall x(U(t,x)\to f_t(A(x))),$$
$$f_t(\Box A) = \forall s(tRs \to f_s(A)).$$

Let G_2 be the first order theory with the axioms

$$\forall x\exists yU(x,y),$$
$$\forall y\forall t\forall s(U(t,y)\to U(s,y)),$$
$$\forall xyz(xRy\&yRz \to xRz).$$

We have $G_0 \vdash \varphi$ (i.e., φ is a theorem of modal K4 with constant domains) iff $G_2 \vdash \forall t f_t(\varphi)$.

Thus the translation f gives us an *implementation* of the modal logic in classical logic. It is important to note at this stage that there may be a variety of implementations. For example modal K4 is complete for Kripke semantics with additional conditions on the accessibility relation such as irreflexivity and tree property. So we can add $\forall x\neg xRx$ to G_2.

Let us now expand modal logic K4 with constant domains with the additional unary predicate $\lambda xW(x)$ and the additional semantic condition

$$t \models W(x) \text{ iff } x = t.$$

Such a predicate W was recommended to be always a part of modal logic (see [Gabbay and Malod, 2002]. It can be axiomatised by the following theory G_1 in the language $\mathcal{L}_1 = \mathcal{L}_0 + \{W\}$:

$$\exists xW(x),$$
$$\forall x(W(x)\to\Box\neg W(x)),$$
$$\forall x(\Diamond(W(x)\&A)\to\Box(W(x)\to A)),$$
$$\forall x\forall y(\Diamond(W(x)\&W(y))\to\Box(A(x)\equiv A(y))).$$

It is known that $(\mathcal{L}_1, G_1)$ is a conservative extension of $(\mathcal{L}_0, G_0)$. However, since interpolation does not hold, we would expect that the modularity theorem does not hold. Let us see what happens to the obvious extension g of the translation f to the predicate W. We need to take $\mathcal{L}_3$ as the language $\mathcal{L}_2$ together with W^* as well. So we have in $\mathcal{L}_3$ the predicate $W^*(x,y)$ reading

$$W^*(x,y) \text{ iff } x = y.$$

So W^* is just equality; it does not add anything to the theory G_2. In fact, since we have the axiom $\forall x(W(x) \to \Box\neg W(x))$, G_3 forces R to be irreflexive. So if we do not include irreflexivity in the implementation, G_3 will not be a conservative extension of G_2.

We would like to take advantage of our comment that this subsection is called "further discussion." We are not exactly clear about the relationship between the computer science notion of "implementation" and the logical notion of "interpretation." Also the exact formulation of the modularisation theorem and its use in computer science will influence, of course, its exact connection with interpolation. We need to further study of what is going on here and we hope to present our results in future.

Consider the constant domain axiom of G_2 of our example, namely

$$\forall x \forall y \forall z(U(x,z) \to U(y,z)).$$

We know that modal K4 does not have interpolation if we have this axiom. It does have interpolation without it. We also know that if we have $\lambda x W(x)$ in the language, we do have interpolation. Thus if modularisation is equivalent to interpolation, this means that without this axiom, W can be implemented in a conservative extension of G_2 (see Figure 3.1) and with this axiom it cannot.

At the moment we cannot pin down exactly where the difference manifests itself.

One thing is clear: there is a strong connection with computer science!

ACKNOWLEDGEMENTS

The authors were supported by EPSRC, grant GR/S78162 entitled *Interpolation and Definability.*

BIBLIOGRAPHY

[Artemov, 2001] S. Artemov, Explicit provability and constructive semantics, *Bull. Symb. Logic*, **7** (2001), 1–36.

[Bergstra *et al.*, 1990] J. A. Bergstra, J. Heering, and P. Klint, Module algebra, *J. Assoc. Comput. Mach.*, **37** (1990), no. 2, 335–372.

[Beth, 1953] W. E. Beth, On Padoa's method in the theory of definitions, *Indag. Math.*, **15** (1953), 330–339.

[Bicarregui *et al.*, 2001] J. Bicarregui and T. Dimitrakos and D. Gabbay and T. Maibaum, Interpolation in practical formal development, *Log. J. IGPL*, **9** (2001), 247–259.

[Chagrov and Zakharyaschev, 1993] A. V. Chagrov and M. V. Zakharyaschev, The undecidability of the disjunction property of propositional logics and other related problems, *J. Symb. Logic*, **58** (1993). 49–82.

[Chagrov and Zakharyaschev, 1997] A. V. Chagrov and M. V. Zakharyaschev, *Modal Logic*, Oxford University Press, Oxford, 1997.

[Craig, 1957] W. Craig, Three uses of Herbrandt-Gentzen theorem in relating model theory and proof theory, *J. Symb. Logic*, **22** (1957), 269–285.

[Czelakowski, 1982] J. Czelakowski, Logical matrices and the amalgamation property, *Stud. Log.*, **41**, (1982), 329–341.

[D'Agostino and Hollenberg, 1998] G. D'Agostino and M. Hollenberg, Uniform interpolation, automata and the modal μ-calculus, In: *Advances in Modal Logic, volume I*, Stanford, California, 1998, pp. 73–84.

[Dimitrakos, 1998] T. Dimitrakos, *Formal support for specification design and implementation*, Imperial College, 1998.

[Dimitrakos and Maibaum, 1997a] Th. Dimitrakos and T. S. E. Maibaum, *Uniformity, interpolation and module specification in a Development Workspace*, Imperial College, 1997.

[Dimitrakos and Maibaum, 1997b] Th. Dimitrakos and T.S.E. Maibaum, *Notes on refinement, interpolation and uniformity*, In: Automated Software Engineering – ASE'97, 12th IEEE International Conference, 1997.

[Ehrich, 1982] H.-D. Ehrich, On the theory of specification, implementation and parametrization of abstract data types, *J. Assoc. Comput. Mach.*, **2**9, (1982).

[Fine, 1979] K. Fine, Failure of the interpolation lemma in quantified modal logic *J. Symb. Logic*, **44** (1979), no. 2, 201–206.

[Fitting, 1983] M. C. Fitting, *Proof Methods for Modal and Intuitionistic Logics*, D. Reidel Publishing Co., Dordrecht, 1983.

[Gabbay, 1971] D. M. Gabbay, Semantic proof of Craig's interpolation theorem for intuitionistic logic and its extensions, Part I, Part II., In: *Logic Colloquium '69*, North-Holland, Amsterdam, 1971, pp. 391–410.

[Gabbay, 1972] D. M. Gabbay, Craig's interpolation theorem for modal logics, *Lect. Notes Math.*, **255** (1972), 111–127.

[Gabbay, 1996] D. M. Gabbay, *Labelled Deductive Systems*, Oxford University Press, 1996.

[Gabbay, 1998] D. M. Gabbay, *Fibring Logics*, Oxford University Press, 1998.

[Gabbay, 2008] W. Carnielli, M. Coniglio, D. Gabbay, P. Gouveia, and C. Sernadas, *Analysis and Synthesis of Logics*, Springer, 2008, 604pp.

[Gabbay and Maksimova, 2005] D. M. Gabbay and L. Maksimova, *Interpolation and Definability: Volume 1, Modal and Intuitionistic Logics*, Oxford University Press, Oxford, 2005.

[Gabbay and Malod, 2002] D. Gabbay and G. Malod, Naming worlds in modal and temporal logic, *J. Log. Lang. Inf.*, **11** (2002), 29–65.

[Gabbay *et al.*, 2003] D. M. Gabbay, A. Kurucz, F. Wolter, and M. Zakharyaschev, *Many-Dimensional Modal Logics: Theory and Applications*, Elsevier, Amsterdam, 2003.

[Gabbay and Ohlbach, 1992] D. M. Gabbay and H. J. Ohlbach, Quantifier elimination in second-order predicate logic, In: *Principles of Knowledge Representation and Reasoning (KR92)*, Morgan Kaufmann, 1992, pp. 425–435.

[Gabbay and Olivetti, 2000] D. M. Gabbay and N. Olivetti, *Goal Directed Algorithmic Proof*, Kluwer, 2000.

[Gabbay and Olivetti, 2005] D. M. Gabbay and N. Olivetti, Interpolation in intuitionistic logic programming, *Log. J. IGPL*, (2005).

[Gabbay and Woods, 2004] D. M. Gabbay and J. Woods, *The Reach of Abduction*, Elsevier, 2004.

[Hoogland, 2000] E. Hoogland, Algebraic Characterisations of Various Beth Definability Properties, *Stud. Log.*, **65** (2000), 91–112.

[Maibaum and Sadler, 1985] T. S. E. Maibaum and M. R. Sadler, Axiomatising specification theory, In: *Recent Trends in Data Type Specification*, Springer-Verlag, Berlin, 1985, pp. 171–177.

[Makkai, 1995] M. Makkai, On Gabbay's proof of the Craig interpolation theorem for intuitionistic predicate logic, *Notre Dame J. Formal Logic*, **36** (1995), 364–381.

[Maksimova, 1976] L. L. Maksimova, Principle of variables separation in propositional logics, *Algebra Logic*, **15** (1976), 168–184.

[Maksimova, 1977] L. L. Maksimova, Craig's theorem in superintuitionistic logics and amalgamated varieties of pseudoboolean algebras, *Algebra Logic*, **16** (1977), 643–681.

[Maksimova, 1979] L. L. Maksimova, Interpolation theorems in modal logics and amalgamable varieties of topoboolean algebras, *Algebra Logic*, **18** (1979), 556–586.

[Maksimova, 1982a] L. L. Maksimova, The Lyndon interpolation theorem in modal logics [in Russian], In: *Mathematical Logic and Theory of Algorithms*, Nauka, Novosibirsk, 1982, pp. 45–55.

[Maksimova, 1982b] L. Maksimova, Relevant principles and formal deducibility, In: *Directions in Relevant Logic*, Kluwer, Dordrecht, 1989, pp. 95–97.

[Maksimova, 1991a] L. L. Maksimova, Failure of interpolation and of the Beth property in temporal logics with the operator "the next", *Sib. Math. J.*, **32** (1991), 109–113.

[Maksimova, 1991b] L. L. Maksimova, The tense logic of "the next" does not have the Beth property, *J. Appl. Non-Class. Log.*, **1** (1991), 73–76.

[Maksimova, 1995] L. Maksimova, On variable separation in modal and superintuitionistic logics, *Stud. Log.*, **55** (1995), 99–112.

[Maksimova, 1999a] L. Maksimova, Superintuitionistic logics and projective Beth property, *Algebra Logic,*, **38** (1999), no. 6, 680–696.

[Maksimova, 1999b] L. Maksimova, Projective Beth property in modal and superintuitionistic logics, *Algebra Logic*, **38** (1999), no. 3, 316–333.

[Maksimova, 1999c] L. Maksimova, Interrelations of algebraic, semantical and logical properties for superintuitionistic and modal logics, In: *Logic, Algebra and Computer Science*, Polish Acad. Sci., Warszawa, 1999, 159–168.

[Maksimova, 2000a] L. Maksimova, Strongly decidable properties of modal and intuitionistic calculi, *Log. J. IGPL*, **8** (2000), 797–819.

[Maksimova, 2000b] L. Maksimova, Intuitionistic logic and implicit definability, *Ann. Pure Appl. Logic*, **105** (2000), 83–102.

[Maksimova, 2001] L. Maksimova, Decidability of the projective Beth property in varieties of Heyting algebras, *Algebra Logic*, **40** (2001), 275–286.

[Maksimova, 2002a] L. Maksimova, Definability and Interpolation in Extensons of Johansson Minimal Logic, In: *Paraconsistency: The logical Way to the Inconsistent (Proceedings of the 2nd World Congress on Paraconsistency, Sao Paolo, Brazil, May 2000)*, Marcel Dekker Inc., New York, 2002, pp. 435–443.

[Maksimova, 2002b] L. Maksimova, Projective Beth's properties in infinite slice extensions of the modal logic K4, In: *Advances in Modal Logic, Volume 3*, World Scientific, 2002, pp. 349–363.

[Maksimova, 2002c] L. Maksimova, Complexity of interpolation and related problems in positive calculi, *J. Symb. Logic*, **67** (2002), 397–408.

[Maksimova, 2003a] L. Maksimova, Complexity of some problems in positive and related calculi, *Theor. Comput. Sci.*, **303** (2003), 171–185.

[Maksimova, 2003b] L. Maksimova, Restricted interpolation in modal logics, In: *Advances in Modal Logic, Volume 4*, King's College Publications, London, 2003, pp. 297–311.

[Maksimova, 2003c] L. Maksimova, Restricted amalgamation and projective Beth property in equational logic, *Algebra Logic*, **42** (2003), 712–726.

[Maksimova, 2004] L. Maksimova, Definability in normal extensions of S4 , *Algebra Logic*, **43** (2004), 387–410.

[Maksimova, to appear] L. Maksimova, Decidable properties of logical calculi and of varieties of algebras, In: *Proceedings of Logic Colloquium 2003*, [To appear]

[Marx, 1995] M. Marx, Interpolation, modularization and knowledge representation, In: *Proceedings of Dutch Conference on AI (NAIC'95)*, 1995.

[McRobbie, 1983] M. A. McRobbie, Interpolation theorems for some first-order distributive-free relevant logics, *J. Symb. Logic*, **48** (1983), 522–523.

[Naruse *et al.*, 1998] H. Naruse and Bayu Surarso and H. Ono, A syntactic approach to Maksimova's principle of variable separation for some substructural logics, *Notre Dame J. Formal Logic*, **39** (1998), 94–113.

[Ono, 2003] H. Ono, Substructural logics and residuated lattices — an Introduction, In: *Trends in Logic: 50 Years of Studia Logics*, Kluwer, 2003, pp. 177–212.

[Pitts, 1983] A. M. Pitts, An application of open maps to categorial logic, *J. Pure Appl. Algebra*, **29** (1983), 313–326.

[Rodenburg and van Galbeek, 1988] P. H. Rodenburg and R. J. van Galbbeek, *An Interpolation Theorem in Equational Logic*, Centre for Mathematics and Computer Science, Amsterdam, The Netherlands, No. CS-R8838, 1988.

[Sain, 1989] I. Sain, Amalgamation, epimorphisms and definability properties in algebraic logic, *Bull. Sect. Logic*, **18** (1989), 72–78.

[Schütte, 1962] K. Schütte, Der Interpolationsatz der intuitionistischen Prädikatenlogik, *Math. Ann.* **148** (1962), 192–200.

[Suzuki, 1990] N.-Y. Suzuki, Some syntactical properties of intermediate predicate logics, *Notre Dame J. Formal Logic*, **31** (1990), 548–559.

[Urquhart, 1984] A. Urquhart, The undecidability of entailment and relevant implication, *J. Symb. Logic*, **49** (1984), 1059–1073.

[Urquhart, 1993] A. Urquhart, Failure of interpolation in relevant logics, *J. Philos. Log.*, **22** (1993), 449–479.

[Veloso, 1993] P. A. S. Veloso, A New, Simpler Proof of the Modularisation Theorem for Logical Specifications, *Bull. IGPL*, **1** (1993), 3–12.

[Wolter, 1997] F. Wolter, A note on interpolation property in tense logic, *J. Philos. Log.*, **36** (1997), 545–551.

[Wronski, 1976] A. Wronski, Remarks on Hallden-completeness of modal and intermediate logics, *Bull. Sect. Logic*, **5** (1976), 126–129.

HANS KAMP, JOSEF VAN GENABITH, UWE REYLE

DISCOURSE REPRESENTATION THEORY

1 INTRODUCTION

Discourse Representation Theory, or DRT, is one of a number of theories of dynamic semantics, which have come upon the scene in the course of the past twenty years. The central concern of these theories is to account for the context dependence of meaning. It is a ubiquitous feature of natural languages that utterances are interpretable only when the interpreter takes account of the contexts in which they are made — utterance meaning depends on context. Moreover, the interaction between context and utterance is reciprocal. Each utterance contributes (via the interpretation which it is given) to the context in which it is made. It modifies the context into a new context, in which this contribution is reflected; and it is this new context which then informs the interpretation of whatever utterance comes next.

The focus on context dependence has led to an important shift in paradigm, away from the "classical" conception of formal semantics which sees semantic theory as primarily concerned with reference and truth and towards a perspective in which the central concept is not that of truth but of information. In this perspective the meaning of a sentence is not its truth conditions but its "information change potential" — its capacity for modifying given contexts or information states into new ones. Theories of dynamic semantics, which have been designed specifically to deal with the two-way interaction between utterance and context, all reflect this change of paradigm. Nevertheless, the connection between information and truth is of paramount importance and they are a crucial ingredient of all dynamic theories. DRT differs from certain other dynamic theories [Groenendijk and Stokhof, 1991; Groenendijk and Stokhof, 1990; Chierchia, 1991; Kohlhase *et al.*, 1996; Eijck and Kamp, 1997] in that the role it attributes to truth is especially prominent - so much so, in fact, that some comparisons between the different types of dynamic theories have gone so far as to qualify DRT as "static". There is some justification for this allegation, but nevertheless DRT contains within it the essence of all that distinguishes dynamic semantics from earlier "static" semantic theories, such as in particular, Montague Grammar, which were exclusively concerned with reference, truth and satisfaction.

Context dependence in natural language is an extraordinarily complex and many-faceted phenomenon. Anaphoric pronouns - pronouns which refer back to something that has been introduced previously in the discourse — represent perhaps the most familiar kind of context dependence; and certainly it is the kind that has been most thoroughly investigated, within linguistics, philosophy and Artificial Intelligence. But it is only one of many,

D. Gabbay and F. Guenthner (eds.),
Handbook of Philosophical Logic, Volume 15, 125–394.
DOI 10.1007/978-94-007-0485-5_3,

and to get a proper perspective on context dependence and its theoretical implications it is important to consider others too. Thus a substantial part of this survey will look at cases of context dependence other than anaphoric pronouns, and at the implications they have for the structure of DRT and of dynamic semantics generally.

We will start, however, with a review of the DR-theoretical treatment of pronominal anaphora, retracing the steps which led to its original form. This will motivate the basic formal version of "classical" DRT, in which the central characteristics of the DR-theoretical approach are easiest to recognize.

This is a handbook of philosophical logic. Thus it seems natural to emphasize the general logical architecture of DRT and its philosophical applications. These will be discussed in sections 2.3 and 3.1. Or decision to give priority to these aspects of the theory has forced us to be silent or very brief on others. Thus our choice of DRT-based treatments of natural language phenomena has been guided by the consideration that those we present here should reveal important logical or philosophical issues. Many of the treatments that can be found in the existing DRT literature have been left out.

We also remain almost entirely silent on the quite extensive work on computer implementations of DRT. As the following sections should make clear, the representational character of DRT renders it especially suitable for implementations - as some computer scientists have put it, the theory can be looked upon as a high level program specification. While we see this amenability to implementation as an important feature of DRT, and as one that also has a clear logical and conceptual importance, the specific problems to which implementation gives rise fall outside the horizons that we consider appropriate for this Handbook. However we will discuss, in Section 3.1, versions of DRT which have been partly inspired by the goal to cast the theory in forms which make its computational properties more transparent and thus facilitate implementation in a large variety of computational environments [Asher, 1993; Bos *et al.*, 1994; Muskens, 1996].

Closely connected with the question of implementation is an issue that becomes unavoidable when semantic analyses are made fully explicit. All natural language semantics is concerned with the question how meaning is determined by syntactic form. Thus every explicit semantic analysis must assume some form of syntactic structure for the natural language expressions with which it is concerned. The choice of syntax is something from which many presentations of DRT have tried to remain aloof — out of the conviction that the specifically DR-theoretical contributions that DRT can make to our understanding of semantic problems are largely independent of the details of syntactic theory and thus should be explained in as syntax-neutral a mode as possible. Nevertheless, the general endeavour of linguistic theory — to arrive at an optimal description of all linguistic properties of natural languages — includes the task of finding the optimal account of its syntax no less than finding an optimal account of its semantics. From this perspective the viability of DRT will depend also on its compatibility with what may come to be recognized, perhaps on largely independent grounds, as the best — or the right — theory of syntax.

At present there are versions of DRT building on many of the leading syntactic frameworks — in particular LFG [Kaplan and Bresnan, 1982], HPSG [Pollard and Sag, 1994], and forms of Categorial Grammar [Steedman, 2001] and of GB [Chomsky, 1981]. As the interface problems posed by these different combinations seem to us to have limited repercussions for the logical and philosophical aspects of DRT, this is a part of the DRT literature which we have also decided to pass over. Here as elsewhere we refer the reader to the bibliography.

2 A DYNAMIC AND REPRESENTATIONAL ACCOUNT OF MEANING

Traditionally, formal approaches to natural language semantics have focused on individual sentences and tried to explicate meaning in terms of truth conditions. Nevertheless it had long been acknowledged that content and context are closely related and in fact strongly determine each other. This is nowhere more evident than in the case of multi-sentence natural language texts and discourses which can constitute highly structured objects with a considerable amount of inter- and intra-sentential cohesion. Much of this cohesion can be traced back to anaphoric properties of natural language expressions, that is their capacity to refer back to (or point forward to) other expressions in the text.[1] Pronominals and tense are but two examples of anaphoric devices — devices whose anaphoric nature was realised many years ago but which, it turned out, were difficult to capture with the machinery available within formal semantics in the 60's and 70's.

When formal semantic approaches were extended to capture inter- and intrasentential anaphoric phenomena, it soon became evident that (i) the narrow conception of meaning in terms of truth conditions has to give way to a more dynamic notion and (ii) the traditional analysis of (NP) anaphora in terms of bound variables and quantificational structures has to be modified. Below we briefly retrace some of the basic and by now often rehearsed[2] arguments.

[1]DRT and other dynamic semantic theories focus on textual anaphora. This is not meant to indicate that deictic and common ground etc. anaphora are in any sense considered less important.

[2]C.f. the introductory sections of [Kamp, 1981a], [Heim, 1982], [Groenendijk and Stokhof, 1991], [Groenendijk and Stokhof, 1990] and textbooks such as [Gamut, 1991] and [Kamp and Reyle, 1993].

DRT is probably still best known for its treatment of the inter- and intrasentential anaphoric relations between (originally singular) indefinite NPs and personal pronouns *s/he, it, him, her, his* and *its*. In this section we will concentrate on this part of the theory and somewhat arbitrarily refer to this part as "core DRT".[3]

2.1 Truth Conditions, Discourse and Interpretation in Context

In predicate logic ([Hodges, 2001]) the following two expressions are truth conditionally equivalent

(1) $\exists x\Phi \Leftrightarrow \neg\forall x\neg\Phi$

If Φ is instantiated to $(delegate(x) \wedge arrive(x))$, then the two formulas are approximate semantic representations of

(2) A delegate arrived.

and

(3) It is not the case that every delegate failed to arrive.

While (2) can be extended into the mini-discourse

(4) A delegate^i arrived. She$_i$ registered.

where anaphoric relationships are indicated by subscripts for anaphors and corresponding superscripts for their antecedent head-words, its truth-conditionally equivalent counterpart (3) does not admit of any such extension:[4]

(5) $*$ It is not the case that every delegate ^{i} failed to arrive. She$_i$ registered.

Truth conditions alone fail to capture the contextual dimension of sentence interpretation. Intuitively (and pre-theoretically) the difference between (2) and (3) (and hence the difference between (4) and (5)) can be accounted for as follows: (2) *updates* the initially available context with an antecedent which can be picked up by anaphoric expressions in subsequent discourse; the truth conditionally equivalent (3) doesn't.

[3]Historically this is somewhat inaccurate since the original motivation for the development of DRT was provided by accounts of temporal anaphora. Here it should also be mentioned that DRT did not come completely "out of the blue". Some of the central concepts were in some form or other already present and/or being developed independently at about the same time as the original formulation of DRT in e.g. the work of [Karttunen, 1976], [Heim, 1982] and [Seuren, 1986].

[4]Here and in what follows the asterisk $*$ in (5) indicates linguistic unacceptability.

It might be presumed that at least simple intersentential anaphora of the type illustrated by the well-formed discourse in (4) could be captured with the machinery provided by traditional Montagovian approaches [Montague, 1973]. On this approach sentence sequencing (i.e. the full stop) is analysed as conjunction and the semantic contribution of the antecedent NP (*a delegate*) "put on ice" and later "quantified-in" into a representation for the conjunction of the first and the second sentence in (4) in which the same variable instantiates the subject positions of the two conjoined clauses:

(6) $(\lambda P.\exists x(delegate(x) \wedge P(x))(\lambda y.arrive(y) \wedge register(y))$

(6) can be reduced to

(7) $\exists x(delegate(x) \wedge arrive(x) \wedge register(x))$

On this account, however, a discourse consisting of n sentences may have to be processed in its entirety with NP meanings on hold before finally quantifying-in can take place. Such an approach fails to capture the *on-line* character of discourse processing by a human interpreter. Worse still, this approach delivers wrong results. Consider (8) and (9):

(8) Exactly one delegate arrived. She registered.

(9) Exactly one delegate arrived and registered.

It is not possible to analyse (8) by treating the full stop between the two sentences as conjunction and then quantifying-in the phrase *exactly one delegate* with logical form $\lambda P \exists x(delegate(x) \wedge P(x) \wedge \forall y([delegate(y) \wedge P(y)] \rightarrow x = y))$. For this associates (8) with the truth conditions of (9), as given in (10), whereas the truth conditions of (8) are rather those of (11). In words, (8) rules out than any other delegates arrived while (9) is compatible with this possibility as long as those further delegates did not register.

(10) $\exists x(delegate(x) \wedge arrive(x) \wedge register(x) \wedge \forall y[(delegate(y) \wedge arrive(y) \wedge register(y)) \rightarrow x = y])$

(11) $\exists x(delegate(x) \wedge arrive(x) \wedge \forall y[(delegate(y) \wedge arrive(y)) \rightarrow x = y] \wedge register(x))$

2.2 Donkey Sentences

Traditionally, indefinite NP's have been translated into logic as predications involving existential quantification with intrasentential anaphors referring back to the indefinites as variables bound by the existential quantifiers. In many cases, this approach delivers the right results. However, puzzles associated with "donkey sentences" (originating in the middle ages and discussed in [Geach, 1962 (Third revised edition: 1980)]) show that indefinites cannot be translated uniformly into existential quantifications and demonstrate the need to revise the traditional quantificational bound variable approach to such NP anaphora.

(12) If Pedro^i owns a donkey^j, he$_i$ likes it$_j$.

(13) Every farmer who owns a donkey^j likes it$_j$.

It is widely agreed that (on at least one prominent reading) the truth conditions associated with (12) and (13) correspond to (14) and (15), respectively:

(14) $\forall x[(donkey(x) \wedge own(pedro, x)) \rightarrow like(pedro, x)]$

(15) $\forall x \forall y[(farmer(x) \wedge donkey(y) \wedge own(x, y)) \rightarrow like(x, y)]$

In (14) the indefinite NP *a donkey* in (12) surfaces as a universally quantified expression taking wide scope over the material implication operator. By contrast, in a sentence like (2) the indefinite *a delegate* has existential import. The occurrence of the indefinite noun phrase *a donkey* in (13) poses similar problems. The indefinite NP, this time located inside a relative clause modifying a universally quantified NP, surfaces as a universally quantified expression with wide scope in (15).

Interpreting (12) under the quantifying-in approach illustrated in (6)–(11) results in

(16) $\exists x(donkey(x) \wedge [own(pedro, x) \rightarrow like(pedro, x)])$

while a direct insertion approach (where quantified NPs are interpreted in situ [Montague, 1973]) produces an open formula, in which the x in the consequent of the material implication is not bound:

(17) $\exists x[donkey(x) \wedge own(pedro, x)] \rightarrow like(pedro, x)$

Neither (16) nor (17) are adequate representations of the perceived meaning (14) of (12). (16) comes out true in case there is (at least) one donkey Pedro doesn't own and (17) doesn't even express a proposition.

2.3 DRT — the Basic Ingredients

Examples (2), (3), (4) and (5) illustrate the need to extend the narrow conception of meaning as truth conditions to a more dynamic notion of meaning relative to context. Examples (8) and (9), (12) and (13) illustrate the need to reconsider the traditional quantificational and bound variable approach to nominal anaphora on the intra- and intersentential level.

In the original formulation of DRT [Kamp, 1981a; Kamp and Reyle, 1993] interpretation involves a two stage process: first, the construction of semantic representations, referred to as Discourse Representation Structures (DRSs), from the input discourse and, second, a model-theoretic interpretation of those DRSs. The dynamic part of meaning resides in how the representations of new pieces of discourse are integrated into the representation of the already processed discourse and what effect this has on the integration of the representations of subsequent, further pieces of discourse. Put differently, a new piece of discourse is interpreted against and in turn updates the representation of the already processed discourse and the meaning of a linguistic expression consists both in its update potential and its truth-conditional import in the resulting representation. The dynamic view of meaning in terms of updates of representations and the attempt at a *rational reconstruction* of the on-line and incremental character of discourse processing by human agents naturally leads to an algorithmic specification of DRS-construction in [Kamp, 1981a; Kamp and Reyle, 1993]. To process a sequence of sentences $S_1, S_2, \ldots, S_n$ the construction algorithm starts with a syntactic analysis of the first sentence S_1 and transforms it in a roughly top-down, left-to-right fashion with the help of DRS construction rules into a DRS K_1 which serves as the context for the processing of the second sentence S_2. The syntactic analysis of S_2 is then added *to* and incrementally decomposed *within* the context DRS K_1. Semantic contributions of constituent parts of S_2 are integrated into DRS K_1 as soon as they become available, eventually resulting in a complete DRS $\mathrm{K}_{1,2}$ for the sequence S_1, S_2. Truth conditional interpretations are provided for completed DRSs K_1, $\mathrm{K}_{1,2}$, $\ldots \mathrm{K}_{1,\ldots,n}$ but not for intermediate steps involving application of DRS construction rules. In its original formulation, DRT tries to do justice to a conception prevalent in a number of AI, Cognitive Science and Linguistics approaches (cf. [Fodor, 1975]) according to which the human mind can be conceived of as an information processing device and that linguistic meanings are best viewed as instructions to dynamically construct and update a mental representation, which can then be employed in further mental processing (such as theoretical and practical reasoning). At the same time, complete meaning representations are associated with truth-conditional semantic interpretations. DRT's decidedly representational stance has inspired (or provoked) research on a large number of "non-representational" approaches to dynamic semantics, cf. [Zeevat, 1989; Groenendijk and Stokhof, 1991; Groenendijk and Stokhof, 1990; Muskens, 1996; Eijck and Kamp, 1997; Harel, 1984].

In the early nineties a new DRT architecture was proposed by Van Der Sandt and Geurts [van der Sandt, 1992; Geurts and van der Sandt, 1999; Geurts, 1999; Kamp, 2001a; Kamp, 2001b], based on a general treatment of presupposition [Soames, 1984]. Informally speaking, a (linguistic) presupposition is a requirement which a sentence imposes on the context in which it is used. If the context doesn't satisfy the presuppositions imposed by the sentence, it may be be modified through "accommodation", i.e. modified or updated to a new context which does satisfy them. If the context neither satisfies all the presuppositions of the sentence nor can be accommodated to one that does, then interpretation aborts; these are cases in which interpreters perceive the sentence as incoherent in the context in which it occurs. Within such an account of presupposition anaphoric expressions such as pronouns can be treated as carrying presuppositions of a special kind, viz. that a suitable antecedent is available for them.

Within the new DRT architecture presuppositions are treated via a two stage procedure. First, a "preliminary" representation is constructed for each individual sentence in which all presuppositons which the sentence carries are given explicit representations. During the second stage the presuppositions represented in the preliminary representation are checked against the context; when necessary and feasible, the context is accommodated. When all presuppositions have been satisfied, the remaining non-presuppositional part of the preliminary representation is merged with the (original or updated) context; the result is a DRS which includes both the context information (possibly with its accommodations) and the contribution made by the sentence.

A further difference between the original version of DRT and the new version is that in the former representations are constructed top-down – the syntactic structure of a sentence is decomposed starting from -the top node which represents the sentence as a whole — whereas in the new version construction proceeds bottom-up: the preliminary representations are constructed from syntactic trees by first assigning semantic representations to the leaves of the tree and then building representations for complex constituents by combining the representations of their immediate syntactic parts. In this section we will give a brief impression of both the old and the new architecture. (Details of the old version of DRT can be found in [Kamp and Reyle, 1993]. For alternative bottom-up constructions see e.g. [Asher, 1993; Muskens, 1996; Eijck and Kamp, 1997].)

We begin with a description of some of the basic tools (such as DRSs, DRS conditions, accessibility, etc.) which are characteristic for the general DRT enterprise.

The DRT-based solutions to interpretation in context and context update (with inter- and intrasentential anaphora) are based on (i) a novel conception of logical form and (ii) the use of Discourse Referents (DRs) to represent the semantic contributions made by noun phrases (as well as certain other sentence elements). The logical forms of DRT are the DRSs already mentioned. DRSs can be extended and merged, and in this way DRSs representing sentences can be combined into DRSs that represent multi-sentence discourses. DRs are DRS constituents which serve to represent entities and which could be described as "variables" that are subject to a novel form of binding. This new form of binding allows among other things for a new treatment of indefinite NPs (which are among the contributors of DRs to sentence and discourse representations), a treatment which accounts for their potential as anaphoric antecedents to pronouns (recall,e.g., the difference between (4) and (5)).

Discourse Representation Structures

Semantic representations in DRT are specified in terms of a language of DRSs. Simple DRSs are pairs consisting of a set of discourse referents U — often referred to as the "universe" of the DRS — and a set of conditions Con. The general form of a DRS is as in (18).

(18) $\langle$ U , Con $\rangle$

Intuitively, the universe collects the discourse entities talked about in a discourse while the conditions express constraints (properties, relations) on those discourse entities. Simplifying somewhat,[5] sentence (2) (*a delegate arrived*) corresponds to the DRS

(19) $\langle$ {x},{delegate(x), arrive(x)} $\rangle$

or, in the often used pictorial "box notation",

(20) $$\begin{array}{|c|} \hline \mathrm{x} \\ \mathrm{delegate(x)} \\ \mathrm{arrive(x)} \\ \hline \end{array}$$

In what follows we will make use of *both* the box notation and the linear notation. The box notation provides better readability especially in the case of complex DRSs (it displays the anaphoric possibilities provided by a context at a glance) while the linear, set based notation saves space and is the basis for formal definitions of syntax, semantics and proof systems for the DRS language.

Informally, the indefinite *a delegate* in (2) contributes the discourse referent x to the universe of the DRS in (20) and the atomic condition delegate(x) to its set of conditions. The VP *arrived* contributes the atomic condition arrive(x). The associated semantics (cf. Definition 10) ensures that this simple DRS is true in a model just in case there *exists* a mapping from the discourse referents of the DRS into the universe of the model such that all the conditions in the set of conditions come out true. In this way discourse referents in the top box of a DRS are endowed with existential force and sets of conditions are interpreted conjunctively.

[5] Abstracting away from tense, aspectual phenomena etc.

DRS Conditions and Accessibility

Discourse referents have a double function. On the one hand they serve as antecedents for anaphoric expressions such as pronouns, on the other they act as the bound variables of quantification theory. This second function entails that discourse referents must be able to stand to each other in certain scope relations. To mark these relations we need the concept of a "sub-DRS": DRSs can occur as constituents of larger DRSs. As it turns out, this mechanism provides a natural explanation of the chameleonic quantificational import (existential or universal) of indefinite NPs like those in (2), (12) or (13). Sub-DRSs always occur as part of *complex* DRS conditions. By contrast, the DRS conditions we have seen so far are *simple* or *atomic* DRS conditions. Two examples of complex DRS conditions are those involving implication and negation.

Conditional sentence constructions of the form *if* S_1 *then* S_2 such as (12) involve a complex DRS condition of the form:

(21) $$\boxed{\mathrm{K}_{\mathrm{S}_1}} \Rightarrow \boxed{\mathrm{K}_{\mathrm{S}_2}}$$

which consists of DRSs for the sentences S_1 and S_2, respectively, joined by the $\Rightarrow$ operator. Similarly, negation introduces a complex condition of the form

(22) $$\neg \boxed{\mathrm{K}_{\mathrm{S}}}$$

which contains the DRS K_{S} for the sentence S in the scope of the negation as its subconstituent. To give an example, sentence (12) gives rise to the following DRS:

(23) $$\begin{array}{|c|}\hline \mathrm{x} \\ \mathrm{predro(x)} \\ \begin{array}{|c|}\hline \mathrm{y} \\ \mathrm{donkey(y)} \\ \mathrm{own(x,y)} \\ \hline \end{array} \Rightarrow \begin{array}{|c|}\hline \mathrm{z}\ \ \mathrm{w} \\ \mathrm{beat(z,w)} \\ \mathrm{z = x} \\ \mathrm{w = y} \\ \hline \end{array} \\ \hline \end{array}$$

The truth conditions (cf. Definition 10) associated with DRS (23) involve a wide scope universal quantification over the discourse referent y associated with the indefinite *a donkey.* Intuitively the interpretation of the conditional says that in order for (23) to be true it must be the case that *whenever* a situation obtains that satisfies the description provided by the antecedent of the conditional, then a situation as described by the consequent obtains as well. In other words, the consequent is *interpreted and evaluated* in the context established by the antecedent. The natural language paraphrase of the truth conditions associated with (23) expresses the universal force with which the indefinite *a donkey* in (12) is endowed. Furthermore, since the consequent is interpreted in the context set by the antecedent, the truth-conditional requirement that situations in which the antecedent is true be accompanied by situations in which the consequent is true is tantamount to situations of the former kind being part of (possibly more comprehensive) situations in which antecedent and consequent are true together. This is the informal justification of why discourse referents introduced in the antecedent of the conditional are available for resolution of anaphors in the consequent but not vice versa. It also explains why the universal quantifier expressed by the conditional is *conservative* in the sense of generalized quantifier theory [Westerstahl, 1989b]. The conservativity of other natural language quantifiers follows in the same way, cf. Section (3.3) below. DRS construction for a universal NP with a relative clause containing an indefinite NP, as in (13), proceeds in a similar manner.

The semantics of conditional DRS conditions, then, is based on the principle that the interpretation of the antecedent can be extended to an interpretation of the consequent. This principle entails that a pronoun in the consequent can be interpreted as anaphoric to a constituent in the antecedent, i.e. the pronoun's discourse referent can be linked to the one introduced by this constituent. Such anaphoric links are subject to what is called *accessibility* in DRT, a relation which must hold between the linked discourse referents and which obtains if, informally speaking, the pronoun occurs within the logical scope of its antecedent. On the other hand, discourse referents from the consequent of a conditional are in general not accessible to pronouns in the antecedent. So there is an asymmetry in the accessibility relation here: discourse referents introduced by constituents in the antecedent are accessible to the consequent but not vice versa (unless they are allowed to "escape" to a higher position in the DRS, cf. the discussion on proper names below). The accessibility relation turns out to play a central role in the DR-theoretical account of when anaphora is possible and when not. How DRS-constructors — which, like those of (21) and (22), create complex DRS conditions — affect accessibility, is an essential aspect of the semantic analysis of the natural language constituents (*if* ... (*then*) ..., *not* etc.) which they are used to represent. It can be argued, along lines similar to the argument we have given for conditionals above, that the discourse referents within the scope of a negation operator $\neg$ are not accessible from outside the SubDRS which is in the scope of the negation operator and similarly for discourse referents in the scope of a conditional operator $\Rightarrow$ (again, unless they can "escape"). As long as $\Rightarrow$ and $\neg$ are the only complex DRS condition constructors, the accessibility relation can be graphically described in terms of the geometrical configurations of the box representation of the DRS language as going *left and up.*

The structure of the DRS determines, via its model-theoretic interpretation, the quantificational import of discourse referents it contains. In this way indefinites are interpreted as terms which receive different quantificational import depending on where the discourse referents they introduce end up within the DRS. To a considerable extent, therefore, the variable binding role of quantifiers in traditional predicate logic or within the higher type Intensional Logic used in Montague Grammar style representations is taken over in DRT by the DRS universes, which in effect act as quantifier prefixes, and by the structure of DRSs which defines the scope and binding properties of these DRS universes.

We are now in a position to account for the contextually relevant difference between the truth-conditionally equivalent (2) and (3) that is manifest in (4) and (5). (2) and (3) are mapped into the DRSs in (24) and (25), respectively:

(24) $$\boxed{\begin{array}{c} \text{x} \\ \text{delegate(x)} \\ \text{arrive(x)} \end{array}}$$

(25) $$\boxed{\neg \boxed{\boxed{\begin{array}{c} \text{x} \\ \text{delegate(x)} \end{array}} \Rightarrow \boxed{\neg \boxed{\text{arrive(x)}}}}}$$

(24) and (25) are truth-conditionally equivalent, as can be verified against the semantics given in Definition 10. However, (24) can be extended to an anaphorically resolved DRS

(26) $$\boxed{\begin{array}{c} \text{x y} \\ \text{delegate(x)} \\ \text{arrive(x)} \\ \text{register(y)} \\ \text{y = x} \end{array}}$$

representing the two sentence discourse (4), while (25) can only be extended to the unresolvable

(27) $$\boxed{\begin{array}{c} \mathrm{y} \\ \neg\ \boxed{\begin{array}{c} \boxed{\begin{array}{c} \mathrm{x} \\ \mathrm{delegate(x)} \end{array}} \Rightarrow \boxed{\neg\ \boxed{\begin{array}{c} \\ \mathrm{arrive(x)} \end{array}}} \end{array}} \\ \mathrm{register(y)} \\ \mathrm{y} = ? \end{array}}$$

where the remaining resolution instruction y = ? indicates that no antecedent for the pronoun *she* has been found.

Finally, let us consider the pair of sentences in (8) and (9). An analysis of sentence sequencing as conjunction together with a quantifying-in approach as the last step in the derivation would ascribe a complex property $\lambda x.(arrive(x) \wedge register(x))$ to the representation $\lambda P.\exists x(delegate(x) \wedge P(x) \wedge \forall y[(delegate(y) \wedge P(x)) \rightarrow x = y])$ of the quantifying NP *exactly one delegate* in one fell swoop resulting in the formula $\exists x(delegate(x) \wedge arrive(x) \wedge register(x) \wedge \forall y([delegate(y) \wedge arrive(y) \wedge register(y)] \rightarrow x = y))$. In contrast, in the DRT approach a discourse referent x is set up by the indefinite NP *a delegate* in the first sentence and then incrementally constrained by the addition of further conditions:

(28) $$\boxed{\begin{array}{c} \mathrm{x\ \ z} \\ \mathrm{delegate(x)} \\ \mathrm{arrive(x)} \\ \boxed{\begin{array}{c} \mathrm{y} \\ \mathrm{delegate(y)} \\ \mathrm{arrive(y)} \end{array}} \Rightarrow \boxed{\begin{array}{c} \\ \mathrm{x = y} \end{array}} \\ \mathrm{register(z)} \\ \mathrm{z = x} \end{array}}$$

In this way we obtain the truth conditions associated with the predicate logic formula $\exists x(delegate(x) \wedge arrive(x) \wedge \forall y[(delegate(y) \wedge arrive(y)) \rightarrow x = y] \wedge register(x))$ which are those intuitively associated with (8) .

DRS Construction

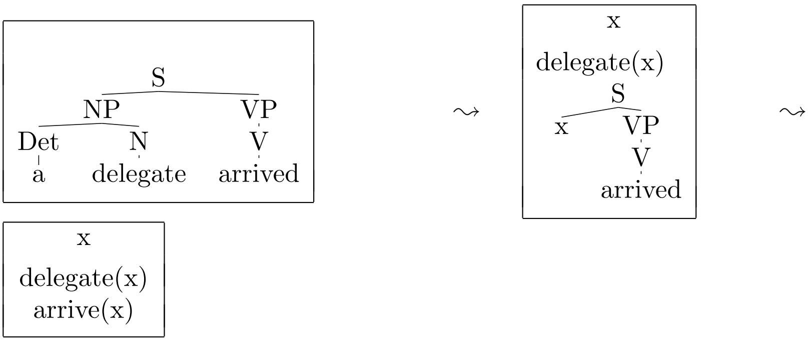
DRS construction has been defined for many of the leading syntactic theories including (simple or decorated) CFG [Kamp, 1981a; Kamp and Reyle, 1993; Bos *et al.*, 1994], LFG [Reyle and Frey, 1983; Genabith and Crouch, 1999], HPSG [Frank and Reyle, 1995] and Categorial Grammar [Zeevat *et al.*, 1987] based approaches. Below we sketch the original top-down DRS construction algorithm and the more recent, bottom-up, presupposition-based version.

A DRT-Top-Down Construction Algorithm: In the original formulation of DRT [Kamp, 1981a; Kamp and Reyle, 1993] the construction of DRSs is spelled out in terms of an algorithm based on DRS construction rules which successively decompose syntactic analyses for the individual sentences in a discourse into DRSs in a roughly top-down, left-to-right manner. Here we briefly and informally illustrate the algorithm with the two sentence mini-discourse (4), here repeated as (29):

(29) A delegate^1 arrived. She$_1$ registered.

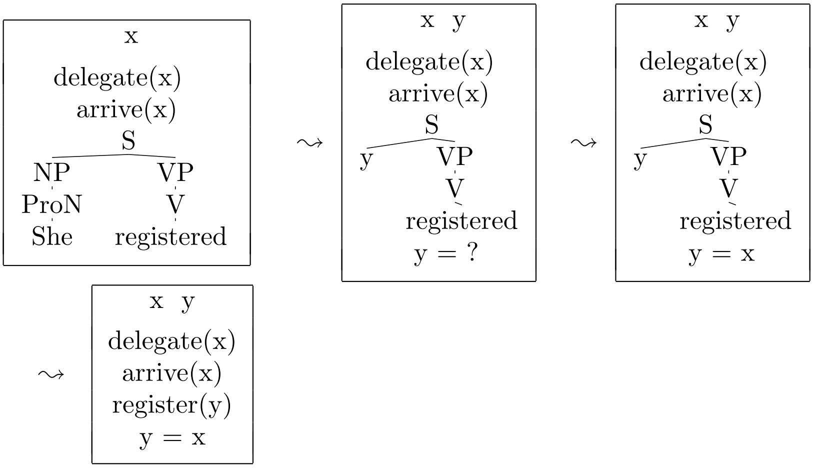
As a first step the algorithm inserts the syntactic analysis of the first sentence in (29) as a "reducible condition" into an empty DRS representing an initial empty context. A DRS construction rule for indefinite NPs matches the relevant part of the tree, introduces a *new* discourse referent x into the universe of the DRS under construction and adds a condition delegate(x) to the set of conditions. The matching part of the tree configuration is replaced by x. Next, a DRS construction rule for simple intransitive VP configurations applies. The matching tree is consumed and a condition arrive(x) is added to the DRS condition. This completes the processing of the first sentence of (29).

(30)



In the next step the top-down construction algorithm inserts the syntactic analysis of the second sentence in (29) as a reducible condition into the context DRS constructed from the first sentence. A DRS construction rule for pronominal NPs introduces a new discourse referent y into the universe of the DRS under construction, adds a condition y = ? to the set of conditions and replaces the matching part of the tree with y. Informally, y = ? can be understood as an instruction to find a suitable antecedent for the pronoun *she*. A suitable antecedent is a discourse referent already introduced and available in the context representation constructed so far. In the case at hand discourse referent x is available and the anaphor is resolved to y = x. Note that in this set-up the anaphoric NP *she* is resolved as soon as it is processed by the construction algorithm. The original DRS construction algorithm was in fact designed as a reconstruction of the on-line and incremental interpretation of a discourse by a human interpreter. In the final step the algorithm processes the intransitive VP in the same fashion as for the first sentence in (29):

(31)



Two Stage Bottom-Up DRS Construction: A top-down DRS construction algorithm of the kind sketched for a modest fragment of English is spelled out in detail in [Kamp and Reyle, 1993]. We already noted that the new, presupposition-based version of DRT makes use of a bottom-up construction process. In recent times bottom-up construction became increasingly common within DRT, and it will be assumed (if often only implicitly) throughout most of this survey. In the next few pages we present, briefly and informally, the essential steps involved in constructing a DRS for the mini-discourse in (29) in the more recent bottom-up and presupposition-based version of DRT.

As noted above, in the new version of DRT DRS construction proceeds in two stages: a preliminary sentence representation is constructed during the first stage and during the second stage the pesuppositions of the sentence, which are explicitly represented in the preliminary DRS, are verified in their respective contexts, with or without context accommodation; when presupposition verification is successful, the non-presuppositional remainder of the preliminary representation is merged with the context representation (or with the representation of the accommodated context). The simplest preliminary representations for sentences with presuppositions are of the form $\langle$ P,D $\rangle$, where D (a DRS) is the non-presuppositional part of the representation and P is a set of representations of the presuppositions of the sentence, where these representations also take the form of DRSs. In more complicated cases the set P may itself consist of preliminary DRSs (as a presupposition may rest in its turn on other presuppositions) and D too may have a more complicated structure which involves additional presuppositions.

We also noted that anaphoric pronouns are treated as carrying a presupposition that the context provides a suitable anaphoric antecedent. In fact, in the new version of DRT all definite NPs are treated as coming with a presupposition to the effect that there is a way of determining their reference which is independent of the remaining material on the sentential utterance to which the NP belongs; reference via coreference with an anaphoric antecedent is one of the various forms which this presupposition can take.

Indefinite NPs, on the other hand, are assumed to be without presupposition. It is this which sets them apart from definite NPs and allows them to make the quantifier-like contributions to sentence meaning which motivated the traditional treatment of indefinites as existential quantifiers. However, the novel form of binding which we mentioned earlier as one of the distinctive features of DRT, and which applies in particular to the discourse referents contributed by indefinites, distinguishes indefinites form "genuine" quantifier phrases like *every delegate* and makes it possible to account for the capacity of indefinites to act as antecedents for anaphoric pronouns in sentences like (12) and (13) and discourses like (4).

The absence of presuppositions connected with indefinites means that no presupposition is introduced by the subject NP *a delegate* of the first sentence of (29). So, if we assume that no other constituent of this sentence carries a presupposition, then the preliminary representation of the sentence will be that given on the right hand side in (32). The left hand side gives the representation of the context, which we have assumed to be empty.

(32) $$\begin{array}{|c|}\hline \quad \\ \quad \\ \hline\end{array} \qquad \left\langle \emptyset\ ,\ \begin{array}{|c|}\hline \text{x} \\ \hline \text{delegate(x)} \\ \text{arrive(x)} \\ \hline\end{array} \right\rangle$$

context preliminary DRS

As there are no presuppositions to resolve, the non-presuppositional part of the preliminary DRS can be merged with the (initially) empty context. Here $\uplus$ is the symmetric merge operation, i.e. $\langle\ U_1,Con_1\ \rangle \uplus \langle\ U_2,Con_2\ \rangle = \langle\ U_1 \cup U_2,Con_1 \cup Con_2\ \rangle$.[6]

[6] There exists an extensive literature on symmetric and non-symmetric merge operations including [Fernando, 1994; Vermeulen, 1995; Eijck and Kamp, 1997].

(33)
$$\begin{array}{|c|}\hline \\ \\ \\ \hline\end{array} \quad \uplus \quad \begin{array}{|c|}\hline \text{x} \\ \text{delegate(x)} \\ \text{arrive(x)} \\ \hline\end{array} \quad = \quad \begin{array}{|c|}\hline \text{x} \\ \text{delegate(x)} \\ \text{arrive(x)} \\ \hline\end{array}$$

context non-presuppositional DRS

The result of the merge in (33) constitutes the new context DRS against which the preliminary DRS for the second sentence in (29) is interpreted:

(34)
$$\begin{array}{|c|}\hline \text{x} \\ \text{delegate(x)} \\ \text{arrive(x)} \\ \hline\end{array} \qquad \left\langle \left\{ \begin{array}{|c|}\hline \text{y} \\ \text{fem(y)} \\ \hline\end{array} \right\}, \begin{array}{|c|}\hline \text{y} \\ \text{register(y)} \\ \hline\end{array} \right\rangle$$

context preliminary DRS

The presuppositional part of the preliminary DRS for the second sentence derives from the pronominal NP *She*. *She* requires a suitable antecedent, either one that has the property female or one that is neutral between a fe/male interpretation, to be available in the context established so far. In (34) a possible antecedent is provided in the form of the discourse referent x for *a delegate* in the context DRS. *Delegate* is neutral (delegates can be either female or male), so the presupposition can be satisfied by accommodating fem(x) to the context DRS. Presupposition resolution is recorded as y = x in the non-presuppositional part of the preliminary DRS and, in the final step, merging the non-presuppositional part of the DRS for the second sentence in (29) with the (updated) context DRS established by the first sentence results in:

(35)
$$\begin{array}{|c|}\hline \text{x} \\ \text{delegate(x)} \\ \text{arrive(x)} \\ \text{fem(x)} \\ \hline\end{array} \quad \uplus \quad \begin{array}{|c|}\hline \text{y} \\ \text{register(y)} \\ \text{y = x} \\ \hline\end{array} \quad = \quad \begin{array}{|c|}\hline \text{x y} \\ \text{delegate(x)} \\ \text{arrive(x)} \\ \text{fem(x)} \\ \text{register(y)} \\ \text{y = x} \\ \hline\end{array}$$

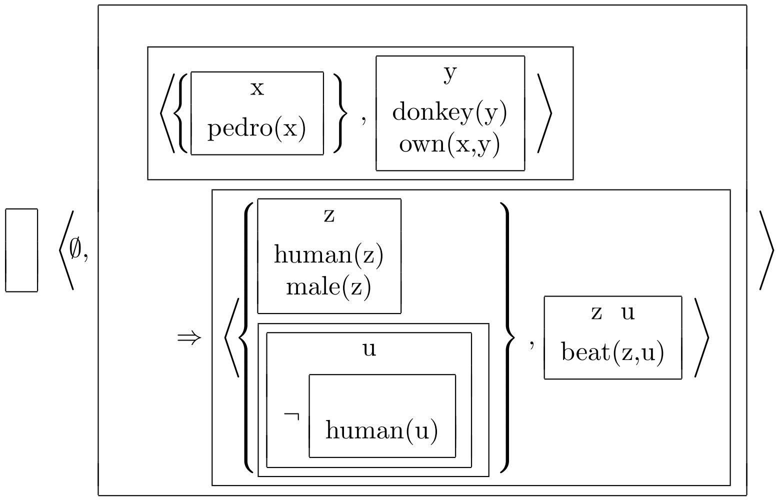
context non-presuppositional DRS

Presupposition verification in (34) involves a "world knowledge" inference corresponding to an axiom of the form $\forall x(delegate(x) \rightarrow (male(x) \vee_x female(x)))$.[7] The example may seem trivial but, in general, presupposition verification may potentially draw upon open-ended knowledge. Except for the accommodated condition fem(x), (35) is equivalent to the final

[7] $\vee_x$ is exclusive or: $P \vee_x Q$ iff $(P \vee Q) \wedge \neg(P \wedge Q)$.

unreducible DRS in (31) obtained via the top-down construction algorithm. Its truth conditions are those of $\exists x \exists y(delegate(x) \wedge arrive(x) \wedge fem(x) \wedge register(y) \wedge y = x)$.

Finally, we outline how example (12) *if Predro owns a donkey, he beats it* comes to be interpreted as the DRS in (23) in the bottom-up approach. The pair consisting of the empty context DRS and the preliminary DRS constructed for (12) is given in (36):

(36) 

(36) is an example of a complex preliminary DRS with embedded presuppositions. The antecedent of the implicational DRS condition contains a presupposition triggered by the proper name *Pedro*. The use of the proper name *Pedro*, as opposed to, say, the phrase *someone named "Pedro"*, carries the implication that Pedro is already part of the available context. To do justice to this intuition, DRT assumes that the discourse referents for proper names are always part of the highest DRS universe (the highest DRS universe contains those discourse referents which represent entities that can be considered as elements of the current context of interpretation, as it has been established by the interpretation of the already processed parts of the text). Note that there is a certain tension between the claim we just made that the use of a name presupposes its bearer to be *already represented in* the context, and the stipulation that the name introduces a discourse referent representing its bearer into the context. This apparent contradiction can be easily resolved. By using the name the speaker presupposes familiarity with it, in the sense of there being a representation of its bearer. This, however, is a type of presupposition that is readily accommodated when neccessary: if the bearer is not yet represented in the context as the interpreter has it, then the context is readily updated by adding a representation for the name's bearer. The discourse referent introduced by the

name's current use can be identified with this representation. Processing the presupposition in the antecedent of the DRS condition in (36) results in the following representation:

(37)
$$\begin{array}{|c|}\hline \text{x} \\ \text{Pedro(x)} \\ \hline \end{array} \left\langle \emptyset, \begin{array}{|c|}\hline \begin{array}{|c|}\hline \left\langle \emptyset, \begin{array}{|c|}\hline \text{y} \\ \text{donkey(y)} \\ \text{own(x,y)} \\ \hline \end{array} \right\rangle \\ \hline \end{array} \\ \Rightarrow \begin{array}{|c|}\hline \left\langle \left\{ \begin{array}{c} \begin{array}{|c|}\hline \text{z} \\ \text{human(z)} \\ \text{male(z)} \\ \hline \end{array} \\ \begin{array}{|c|}\hline \text{u} \\ \neg \begin{array}{|c|}\hline \\ \text{human(u)} \\ \hline \end{array} \\ \hline \end{array} \end{array} \right\}, \begin{array}{|c|}\hline \text{z u} \\ \text{beat(z,u)} \\ \hline \end{array} \right\rangle \\ \hline \end{array} \\ \hline \end{array} \right\rangle$$

The consequent of the preliminary DRS involves two presuppositions generated by the pronominal NPs *he* and *it*. The former requires an antecedent that is human and male, the latter an antecedent that is nonhuman. Pronominal presuppositions can only be resolved through satisfaction by the local or nonlocal context, which is to say that the required antecedents will have to be provided by the context. The context available to both pronominals is provided by (i) the antecedent of the implicational condition, together with (ii) the context DRS and (iii) the discourse referents and conditions of the DRS which contains the ⇒-condition as a component. (However, in the present case in which the universe of this DRS is empty, this third component of the context has no part to play.) These are precisely the domains that are accessible from the position of the consequent of the conditional, in the sense of accessibility alluded to in the description of the top-down algorithm. Also, the antecedent-presupposition triggered by the name *Pedro* has already been accommodated and the result of this accommodation added to the (previously empty) context DRS. Thus the presupposition introduced by *he* can be resolved at the level of the context DRS while the one for *it* is resolved at the level of the antecedent DRS. These resolutions match *he* with the discourse referent x introduced by *Pedro* and *it* with discourse referent y introduced by *a donkey.* These matches are recorded by z = x and u = y, which are added to the non-presuppositional component of the consequent DRS. Again, presupposition verification involves "world-knowledge". The resulting representation is[8]

[8]Here and elsewhere we sometimes supress (some of) the presuppositional constraints, such as e.g.

$$(38)\quad \boxed{\begin{array}{c} \mathrm{x} \\ \mathrm{pedro(x)} \end{array}} \left\langle \emptyset, \boxed{ \boxed{ \left\langle \emptyset, \boxed{\begin{array}{c} \mathrm{y} \\ \mathrm{donkey(y)} \\ \mathrm{own(x,y)} \end{array}} \right\rangle } \Rightarrow \boxed{ \left\langle \emptyset, \boxed{\begin{array}{c} \mathrm{z\ u} \\ \mathrm{beat(z,u)} \\ \mathrm{z{=}x} \\ \mathrm{u{=}y} \end{array}} \right\rangle } } \right\rangle$$

At this stage all presuppositions generated by the preliminary DRS are resolved (or cancelled) and the various (local and global) context DRSs can now be merged with the non-presuppositional components of the preliminary DRS to yield the DRS for the discourse in (12):

$$(39)\quad \boxed{\begin{array}{c} \mathrm{x} \\ \mathrm{predro(x)} \end{array}} \uplus \boxed{ \boxed{\begin{array}{c} \mathrm{y} \\ \mathrm{donkey(y)} \\ \mathrm{own(x,y)} \end{array}} \Rightarrow \boxed{\begin{array}{c} \mathrm{z\ \ u} \\ \mathrm{beat(z,u)} \\ \mathrm{z = x} \\ \mathrm{u = y} \end{array}} } \quad =$$

$$\boxed{\begin{array}{c} \mathrm{x} \\ \mathrm{pedro(x)} \\ \boxed{\begin{array}{c} \mathrm{y} \\ \mathrm{donkey(y)} \\ \mathrm{own(x,y)} \end{array}} \Rightarrow \boxed{\begin{array}{c} \mathrm{z\ \ u} \\ \mathrm{beat(z,u)} \\ \mathrm{z = x} \\ \mathrm{u = y} \end{array}} \end{array}}$$

3 BASIC DRS LANGUAGES AND THEIR INTERPRETATIONS

In this section we provide formal definitions of the syntax and semantics of some basic DRS-languages. We start with a simple, extensional, first-order DRS language, present an intensional model for the language and define the notion of proposition expressed by a DRS as the set of possible worlds where the DRS is true. Truth conditions and propositions, however, do not fully capture the dynamic aspects of discourse interpretation in DRT where sentences are interpreted against a previously established context and where, in turn, a given context is updated through this interpretation into a

---

$$\boxed{\begin{array}{c} \mathrm{u} \\ \neg \boxed{\mathrm{human(u)}} \end{array}}$$

to avoid clutter in the resolved representations.

new context for subsequent sentences. We model this dynamics semantically in terms of information states and, based on this, context change potentials (CCPs [Heim, 1982]), i.e. functions on, or relations between information states. The remaining parts of this section considers extensions that deal with generalized quantifiers, plurals, tense and aspect.

3.1 A First-Order, Extensional DRS Language

Here we provide the syntax and semantics for simple, complete and proper DRSs, the final products resulting from exhaustive and successful application of either of the two DRS construction algorithms informally presented in Section (2.3). Such DRSs do not contain any reducible conditions (i.e. they are complete) or presuppositions (they are simple) and all occurrences of discourse referents are bound (they are proper).

The vocabularies of simple, first-order, extensional DRS languages consist of four disjoint sets.

DEFINITION 1. The vocabulary for a simple, extensional DRS language L is given by:

(i) a set Ref of discourse referents

(ii) a set Name of one-place definite relation constants

(iii) sets Rel^n of predicate constants

(vi) a set Sym of logical symbols; for the language defined in this section this is the set $\{=, \neg, \vee, \Rightarrow\}$

In languages of this form, the work of individual constants in ordinary predicate logic is done by the unary predicates in the set Name. Thus, instead of an individual constant p to denote *Pedro*, Name will contain a unary predicate Pedro and the condition "Pedro(x)" expresses that x represents the individual *Pedro*.

DRSs and DRS-conditions are defined by simultaneous recursion:

DEFINITION 2. Syntax of DRSs and DRS conditions of L:

(i) if $\mathrm{U} \subseteq \mathrm{Ref}$ and Con a (possibly empty) set of conditions, then $\langle$ U,Con $\rangle$ is a DRS

(ii) if x_i, $\mathrm{x}_j \in \mathrm{Ref}$, then $\mathrm{x}_i = \mathrm{x}_j$ is a condition

(iii) if $\mathrm{N} \in$ Name and $\mathrm{x} \in$ Ref, then N(x) is a condition

(iv) if P is a n-place predicate constant in Rel and $\mathrm{x}_1, \ldots, \mathrm{x}_n \in$ Ref, then $\mathrm{P}(\mathrm{x}_1, \ldots, \mathrm{x}_n)$ is a condition

(v) if K is a DRS, then $\neg$ K is a condition

(vi) if K_1 and K_2 are DRSs, then $K_1 \vee K_2$ is a condition

(vii) if K_1 and K_2 are DRSs, then $K_1 \Rightarrow K_2$ is a condition

The conditions specified in (ii), (iii) and (iv) are called *atomic* conditions, those specified in (v), (vi) and (vii) *complex* conditions .
Given a DRS K, FV(K) denotes the set of *free discourse referents* of K.

DEFINITION 3. FV(K), the set of free discourse referents of K, is defined by:

(i) $FV(\langle U_K, Con_K \rangle) := (\bigcup_{\gamma \in Con_K} FV(\gamma))$ - U_K

(ii) $FV(x_i = x_j) := \{x_i, x_j\}$

(iii) $FV(P(x_1,...,x_n)) := \{x_1, \ldots, x_n\}$

(iv) $FV(\neg K) := FV(K)$

(v) $FV((K_1 \vee K_2)) := FV(K_1) \cup FV(K_2)$

(vii) $FV(K_1 \Rightarrow K_2) := FV(K_1) \cup (FV(K_2)$ - $U_{K_1})$

BV(K), the set of *bound discourse referents* of K, is the set $V(K) \setminus FV(K)$, where V(K) is the set of all discourse referents occurring somewhere in K.

A *proper* DRS is a DRS where all occurrences of discourse referents are properly bound.

DEFINITION 4. A DRS K is proper iff $FV(K) = \emptyset$

To define the notion of a pure DRS formally we need to make use of the relation of one DRS being a sub-DRS of another DRS. This relation, which we represent as $\leq$, is defined as the reflexive transitive closure of the relation $<$ of a DRS K_1 being an immediate sub-DRS of a DRS K. $<$ is given in Definition 5.

DEFINITION 5. K_1 is an immediate sub-DRS of K, $K_1 < K$, if any of the following conditions holds:

(i) $\neg K_1 \in Con_K$

(ii) there is a DRS K_2 sth. $K_1 \Rightarrow K_2 \in Con_K$ or $K_2 \Rightarrow K_1 \in Con_K$

(iii) there is a DRS K_2 sth. $K_1 \vee K_2 \in Con_K$ or $K_2 \vee K_1 \in Con_K$

Purety of a DRS can now be defined as in Definition 6. A DRS is *pure* if it does not contain otiose declarations of discourse referents.

DEFINITION 6. A DRS K is pure iff for every two distinct DRSs K_1 and K_2 such that K_2 is a sub-DRS of K_1 and K_1 is a sub-DRS of K, $U_{K_2} \cap (U_{K_1} \cup FV(K)) = \emptyset$.

On the basis of the relation $\leq$ we can also define a relation of accessibility, either between DRSs of between discourse referents. The accessibility relation between DRSs is given in Definition (7), that between between discourse referents in Definition 8.

DEFINITION 7. Given DRSs K and K_1, K is accessible from K_1, in symbols K *acc* K_1, iff

(i) $K_1 \leq K$; or

(ii) there exist DRSs K_2 and K_3, sth. $K_2 \Rightarrow K_3$ and K *acc* K_2 and K_3 *acc* K_1

DEFINITION 8. Given DRSs K, K_1 and discourse referents x and y, x is accessible from y, in symbols x *acc* y, iff $x \in U_K$, $y \in U_{K_1}$ and K *acc* K_1.

Models $\langle U, \Im \rangle$ for the simple DRS language L defined above are extensional first-order models consisting of a non-empty domain U of individuals and an interpretation function $\Im$ which maps names in Name into elements in U, and n-ary relations in Rel into sets of n-tuples of elements of U, i.e. into elements of the set $\mathcal{P}(U^n)$.

DEFINITION 9. Interpretation functions $\Im$ for models of L are defined as follows:

(i) $\Im : \text{Name} \rightarrow \{\{u\}|u \in U\}$

(ii) $\Im : \text{Rel}^n \rightarrow \mathcal{P}(U^n)$

The model-theoretic interpretation of the core DRS language defined above can be illustrated as follows: by way of a first approximation, a DRS $K = \langle U_K, Con_K \rangle$ can be thought of as a "partial" model (this is qualified below) representing the information conveyed by some discourse D; K is true if and only if K can be embedded into the "total" model $\mathcal{M} = \langle U, \Im \rangle$ by mapping all the discourse referents in the universe U_K of K into elements in the domain U of $\mathcal{M}$ in such a way that under this mapping all the conditions $\gamma \in Con_K$ come out true in $\mathcal{M}$. In other words, K is true if and only if there is a homomorphism from K into $\mathcal{M}$. In DRT parlance, such a homomorphism is called a *verifying embedding for* K *into* $\mathcal{M}$. Embeddings are *partial* variable or discourse referent[9] assignment functions and the notation $g \subseteq_X k$, where

[9]We often use the terms "variable" and "discourse referent" interchangeably.

X is a (possibly empty) set of discourse referents, states that embedding k extends g to the discourse referents in X, i.e. Dom(k) = Dom(g) $\cup$ X.[10]

The conception of a DRS K as a partial model makes straightforward sense only in those cases where all conditions of K are atomic. As soon as the DRS contains complex conditions, of the form (21), say, or of the form (22), the notion becomes problematic for the very same reasons that negation and implication are problematic in Situation Semantics ([Cooper *et al.*, 1990; Barwise *et al.*, 1991; Barwise and Cooper, 1993]). Take negation: should the condition

$$(40) \quad \neg \boxed{\begin{array}{c} \text{y} \\ \hline \text{donkey(y)} \\ \text{own(x,y)} \end{array}}$$

be understood as giving partial information in the sense that (the value of) x does not own any of the donkeys that can be found in some limited set or should it be taken as an absolute denial that x owns any donkeys whatever? The view adopted by classical DRT is that (40) is to be interpreted absolutely, in the sense that an embedding (assignment) f with f(x)=a into a model $\mathcal{M} = \langle$ U , $\Im$ $\rangle$ verifies (40) iff there is no b $\in$ U such that b $\in$ $\Im$(donkey) and $\langle$ a,b $\rangle \in \Im$(own); or to put it into slightly different terms, and assuming that f is not defined for y: f verifies (40) in $\mathcal{M}$ iff there is *no* function g such that $f \subset_{\{y\}} g$ (i.e. no extension g of f such that Dom(g) = Dom(f) $\cup$ {y}) which verifies (41) in $\mathcal{M}$.

$$(41) \quad \boxed{\begin{array}{c} \text{y} \\ \hline \text{donkey(y)} \\ \text{own(x,y)} \end{array}}$$

A similar verification clause is adopted for complex conditions of the form $K_1 \Rightarrow K_2$, where $K_1 = \langle U_{K_1}, Con_{K_1} \rangle$ and $K_2 = \langle U_{K_2}, Con_{K_2} \rangle$ are DRSs. $K_1 \Rightarrow K_2$ is verified by f in $\mathcal{M}$ iff for *every* g such that $f \subset_{U_{K_1}} g$ which verifies the conditions in K_1 there *exists* an h such that $g \subset_{U_{K_2}} h$ and h verifies the conditions in K_2. Putting these considerations together we come to the following definitions of verification and truth:

DEFINITION 10. Verifying embeddings for DRSs and DRS conditions of L:

(i) $\langle g, h\rangle \models_{\mathcal{M}} \langle$ U,Con $\rangle$ iff $g \subset_{U} h$ and for *all* $\gamma \in$ Con: $h \models_{\mathcal{M}} \gamma$

(ii) $g \models_{\mathcal{M}} x_i = x_j$ iff $g(x_i) = g(x_j)$

[10] Below we use $g \subseteq_X k$ and $k \supseteq_X g$ interchangeably.

(iii) $g \models_{\mathcal{M}} \mathrm{N(x)}$ iff $\Im(\mathrm{N}) = \{g(\mathrm{x})\}$

(iv) $g \models_{\mathcal{M}} \mathrm{P(x_1,..,x_n)}$ iff $\langle g(\mathrm{x_1}), .., g(\mathrm{x}_n)\rangle \in \Im(\mathrm{P})$

(v) $g \models_{\mathcal{M}} \neg\, \mathrm{K}$ iff there does *not exist* an h such that $\langle g, h\rangle \models_{\mathcal{M}} \mathrm{K}$

(vi) $g \models_{\mathcal{M}} \mathrm{K_1} \vee \mathrm{K_2}$ iff there is *some* h such that $\langle g, h\rangle \models_{\mathcal{M}} \mathrm{K_1}$ or there is *some* h such that $\langle g, h\rangle \models_{\mathcal{M}} \mathrm{K_2}$

(vii) $g \models_{\mathcal{M}} \mathrm{K_1} \Rightarrow \mathrm{K_2}$ iff *for all* m such that $\langle g, m\rangle \models_{\mathcal{M}} \mathrm{K_1}$ *there exists* a k such that $\langle m, k\rangle \models_{\mathcal{M}} \mathrm{K_2}$

When $g \models_{\mathcal{M}} \gamma$ where γ is a DRS condition, we say that g *verifies* γ in $\mathcal{M}$. When K is a DRS and $\langle g, h\rangle \models_{\mathcal{M}} \mathrm{K}$, we say that h *verifies* K *with respect to* g.

DEFINITION 11. Truth of a proper DRS K in a model $\mathcal{M}$:

> A proper DRS K is true in a model $\mathcal{M}$ iff
> there *exists* a verifying embedding h for K in $\mathcal{M}$ with respect to the empty assignment Λ.
> We write: $\models_{\mathcal{M}} \mathrm{K}$ iff there exists an h such that $\langle \Lambda, h\rangle \models_{\mathcal{M}} \mathrm{K}$.

The definition of truth for a DRS in a model given in 11, together with the definition of a verifying embedding for DRSs in 10, ensures that the discourse referents in the universe of a main DRS (i.e. one which is not occurring as a sub-DRS of some other DRS) are interpreted as existentially quantified variables. The existential quantifier in the truth definition in 11 is often referred to as existential closure. Note the difference between the existential closure which the truth definition imposes on the discourse referents in the universe of a main DRS and the universal quantification imposed on the discourse referents in the antecedent of a conditional DRS condition $\mathrm{K_1} \Rightarrow \mathrm{K_2}$, as shown in clause 10(vii). Note also the conjunctive interpretation that 10(i) imposes on condition sets: in order that h verifies $\langle$ U , Con $\rangle$ (with respect to a prior embedding g) in $\mathcal{M}$, h must extend g to U and h must verify each of the conditions $\gamma_1, \ldots, \gamma_n \in$ Con (which is equivalent to the claim that h verifies their conjunction). Thus it is an effect of 10(i) that conjunction is built into the structure of a DRS via its condition set, just as it follows from 11 that existential quantification is built into it via its universe. There is no need to represent the conjunction and existential quantification operators of classical logic by means of special devices (i.e. in the form of special complex conditions — but see the discussion of dynamic conjunction in Section 4). One consequence of this is that the DRS language in which the only complex conditions are of the form $\neg$ K has the expressive power of the full predicate calculus (for this sub-language can express $\exists$, $\wedge$

and $\neg$, and the other logical operators of classical logic can be expressed with the help of these, cf. [Kamp and Reyle, 1991]).

The DRSs in the first order fragment as defined in 2 and 10 can be mapped straightforwardly into corresponding FOPL formulae in terms of a function $\wp\ell$[11] following the clauses in the syntactic definition 2 above:

DEFINITION 12. Translation of L into FOPL

(i) $\wp\ell(\langle\{\mathrm{x}_1,\ldots,\mathrm{x}_n\},\{\gamma_1,\ldots,\gamma_m\}\rangle) := \exists\mathrm{x}_1\ldots\exists\mathrm{x}_n(\wp\ell(\gamma_1)\wedge\ldots\wedge\wp\ell(\gamma_m))$

(ii) $\wp\ell(\mathrm{x}_i = \mathrm{x}_j) := (\mathrm{x}_i = \mathrm{x}_j)$

(iii) $\wp\ell(\mathrm{N}(\mathrm{x})) := (\mathrm{N} = x)$[12]

(iv) $\wp\ell(\mathrm{P}(\mathrm{t}_1,\ldots,\mathrm{t}_n)) := \mathrm{P}(\mathrm{t}_1,\ldots,\mathrm{t}_n)$

(v) $\wp\ell(\neg\mathrm{K}) := \neg(\wp\ell(\mathrm{K}))$

(vi) $\wp\ell(\mathrm{K}_1 \vee \mathrm{K}_2) := \wp\ell(\mathrm{K}_1) \vee \wp\ell(\mathrm{K}_2)$

(vii) $\wp\ell(\langle\{\mathrm{x}_1,\ldots,\mathrm{x}_n\},\{\gamma_1,\ldots,\gamma_m\}\rangle \Rightarrow \mathrm{K}_2) := \forall\mathrm{x}_1\ldots\forall\mathrm{x}_n[(\wp\ell(\gamma_1)\wedge\ldots\wedge\wp\ell(\gamma_m)) \rightarrow \wp\ell(\mathrm{K}_2)]$

The definition of verifying embeddings for DRSs given in 10 can be regarded as the definition of a relation between partial input and output assignments on discourse referents — the relation which holds between an output assignment o and an input assignment i relative to a model $\mathcal{M}$ and DRS K if o extends i and verifies the conditions of K in $\mathcal{M}$. In the light of this, the input assignment i may be seen as potentially verifying K in $\mathcal{M}$ if it has one or more extensions o verifying the conditions of K in $\mathcal{M}$. Alternatively, verification may be seen as a non-deterministic process which transforms i into one of the possible output assignments o. The input-output view of the verification definition for DRT is very natural from the perspective of the semantics of programming languages [Harel, 1984]. This analogy has led to versions of the semantics of DRT which are very compact and elegant (see e.g. [Dekker, 1993; Muskens, 1996; Kohlhase *et al.*, 1996; Muskens *et al.*, 1997; Eijck and Kamp, 1997]) and inspired alternative approaches (such as [Groenendijk and Stokhof, 1991; Groenendijk and Stokhof, 1990]). In these versions, DRSs are interpreted as programmes consisting of sequences of instructions; some of these take the form of the introduction of a discourse referent, others the form of DRS

[11] Strictly speaking in order to ensure that $\wp\ell$ is functional we have to define it for a certain canonical order on the sets of discourse referents and conditions in a given DRS. The definition given above maps a DRS into a set of equivalent FOPL formulae.

[12] N is assumed to be a first-order predicate logic constant in the right hand side of (iii) denoting the object u in the interpretation $\{u\}$ assigned to the corresponding DRT definite relation symbol on the left hand side.

conditions. A DRS of the form $\langle$U,Con$\rangle$ is one where all the instructions of the first type precede those of the second; but in this new version of DRT, any order of discourse referents and conditions is admissible (though DRSs which differ in the order of their instructions will, in general, not be equivalent, even if they involve the same set of instructions). Here we give a (star free)[13] fragment of Quantificational Dynamic Logic (QDL, cf. [Pratt, 1976; Harel, 1984; Goldblatt, 1992 (first edition 1987)]) and show how simple first-order DRSs can be translated into QDL programmes. QDL standardly assumes total assignments so there is a prima facie mismatch between that semantics and the partial assignment semantics in DRT. However, so long as we restrict attention to pure DRSs the partial assignment semantics can be restated without difficulty as a semantics involving total assignments. The translation given below in Definition 15 preserves satisfaction. Embeddings are also possible if both QDL and DRT are defined with partial assignments, for details see e.g. [Fernando, 1992]. The syntax of (a fragment of) QDL formulas $\mathcal{F}$ and programmes $\mathcal{P}$ is defined by simultaneous recursion:

DEFINITION 13. A QDL Syntax Fragment:

(i) $\mathrm{P}(\mathrm{t}_1, \ldots, \mathrm{t}_n) \in \mathcal{F}$

(ii) $\perp \in \mathcal{F}$

(iii) if $\pi \in \mathcal{P}$ and $\phi \in \mathcal{F}$ then $[\pi]\phi \in \mathcal{F}$

(iv) $\mathrm{x} := \,? \in \mathcal{P}$

(v) if $\pi_1, \pi_2 \in \mathcal{P}$ then $\pi_1; \pi_2 \in \mathcal{P}$

(vi) if $\phi \in \mathcal{F}$ then $\phi? \in \mathcal{P}$

Intuitively, x := ? is a random assignment; $\pi_1; \pi_2$ is a sequence of programmes: first carry out π_1, then π_2. The postfix operator '?' in (13vi) turns formulas into programmes. $[\pi]\phi$ is a formula stating that ϕ will be true after every terminating execution of π. The semantics of QDL is given in terms of ordinary first order models $\mathcal{M} = \langle \mathrm{U}, \Im \rangle$ and total assignment functions $g, i, o, \ldots$:

DEFINITION 14. QDL Semantics

(i) $[\![\mathrm{P}(\mathrm{t}_1, \ldots, \mathrm{t}_n)]\!] = \{g | \langle [\![\mathrm{t}_1]\!]^g, \ldots, [\![\mathrm{t}_n]\!]^g \rangle \in \Im(\mathrm{P})\}$

(ii) $[\![\perp]\!] = \emptyset$

(iii) $[\![[\pi]\phi]\!] = \{g | \text{for all } m \text{ sth. } \langle g, m \rangle \in [\![\pi]\!] \text{ there exists } h \text{ sth. } \langle m, h \rangle \in [\![\phi]\!]\}$

[13] "$*$" is the iteration operator.

(iv) $[\![\mathrm{x} := ?]\!] = \{\langle i, o\rangle | i[\mathrm{x}]o\}$[14]

(v) $[\![\phi_1; \phi_2]\!] = \{\langle i, o\rangle |$there exists an m sth. $\langle i, m\rangle \in [\![\phi_1]\!]$ and $\langle m, o\rangle \in [\![\phi_2]\!]\}$

(vi) $[\![\phi?]\!] = \{\langle i, i\rangle | i \in [\![\phi]\!]\}$

The execution of a programme may change an input state into possibly different output states. States are modelled as sets of embeddings (sets of assignments of values to variables). At a given state a formula is either true or false. The '?' post-fix operator turns a formula into a test, i.e. a programme that passes on the input assignment unchanged if the assignment supports the formula in the scope of the operator; otherwise execution aborts. It is easy to see how negation $\neg\phi$ can be modelled as $[\phi?]\bot$ and existential quantification $\exists x\phi$ as $\langle x =?\rangle\phi$ where $\langle\pi\rangle\phi$ is shorthand for $\neg([\pi](\neg\phi))$. The embedding Q of pure DRSs (Definition 6) into QDL translates DRS conditions into formulas and DRSs into programmes as follows:

DEFINITION 15. DRT to QDL translation:

(i) $\mathrm{Q}(\mathrm{P}(\mathrm{t}_1, \ldots, \mathrm{t}_n)) = \mathrm{P}(\mathrm{t}_1, \ldots, \mathrm{t}_n)$

(ii) $\mathrm{Q}(\neg\, \mathrm{K}) = [\mathrm{Q}(\mathrm{K})]\bot$

(iii) $\mathrm{Q}(\mathrm{K}_1 \Rightarrow \mathrm{K}_2) = [\mathrm{Q}(\mathrm{K}_1)]\langle\mathrm{Q}(\mathrm{K}_2)\rangle\top$

(iv) $\mathrm{Q}([\mathrm{x}_1, \ldots, \mathrm{x}_n — \gamma_1, \ldots, \gamma_m]) = \mathrm{x}_1 = ?\ ; \ldots, \mathrm{x}_n = ?\ ; \mathrm{Q}(\gamma_1)?\ ; \ldots ; \mathrm{Q}(\gamma_m)?$

Working again with partial assignments (embeddings), a discourse referent x is interpreted as an instruction to extend the current assignment (the input assignment) randomly with an assignment to x while the occurrence of a condition γ functions as a check whether an assignment satisfies the constraint γ expresses. If we stick to the DRS format adopted here (as in [Kamp, 1981a] and [Kamp and Reyle, 1993]) — DRSs are pairs $\langle\mathrm{U,Con}\rangle$ — then the input-output perspective can be brought out more prominently in the following reformulation 16 of 10:

DEFINITION 16.

(o) $[\![\mathrm{x}_l]\!]^i = i(\mathrm{x})$ if $\mathrm{x} \in \mathrm{Dom}(i)$; undefined otherwise.

(i) $[\![\langle\mathrm{U}, \mathrm{Con}\rangle]\!] := \{\langle i, o\rangle | i \subset_{\mathrm{U}} o$ and $o \in \bigcap_{\gamma_j \in \mathrm{Con}} [\![\gamma_j]\!]\}$

(ii) $[\![\mathrm{x}_l = \mathrm{x}_k]\!] := \{i | [\![\mathrm{x}_l]\!]^i \text{and} [\![\mathrm{x}_k]\!]^i$ defined and $= [\![\mathrm{x}_l]\!]^i = [\![\mathrm{x}_k]\!]^i\}$

[14]Given two variable assignment functions i and o, $i[\mathrm{x}]o$ states that o is exactly like i except possibly for the value assigned to x.

(iii) $[\![\mathrm{N(x)}]\!] := \{i | [\![\mathrm{x}]\!]^i \text{ defined and } [\![\mathrm{x}]\!]^i \in \Im(\mathrm{N})\}$

(vi) $[\![\mathrm{P}(\mathrm{x}_1, \ldots, \mathrm{x}_n)]\!] := \{i | [\![\mathrm{x}_j]\!]^l \text{ defined for } j = 1, ..., n \text{ and } \langle [\![\mathrm{x}_1]\!]^i, \ldots, [\![\mathrm{x}_n]\!]^i \rangle \in \Im(\mathrm{P})\}$

(v) $[\![\neg\mathrm{K}]\!] := \{i | \neg\exists o \langle i, o\rangle \in [\![\mathrm{K}]\!]\}$

(vi) $[\![\mathrm{K}_1 \vee \mathrm{K}_2]\!] := \{i | \exists o(\langle i, o\rangle \in [\![\mathrm{K}_1]\!]) \text{ or } \exists o(\langle i, o\rangle \in [\![\mathrm{K}_2]\!])\}$

(vii) $[\![\mathrm{K}_1 \Rightarrow \mathrm{K}_2]\!] := \{i | \forall o(\langle i, o\rangle \in [\![\mathrm{K}_1]\!] \rightarrow \exists k \langle o, k\rangle \in [\![\mathrm{K}_2]\!])\}$

In 16 DRS conditions are interpreted as sets of assignments. In other words, they are "externally static" and do not pass on updated assignments to other conditions. Conditions act as tests on the current assignment and pass on the assignment unchanged if it verifies the condition. One condition, $\mathrm{K}_1 \Rightarrow \mathrm{K}_2$, is "internally dynamic". The (possibly updated) output assignments of the antecedent of implicative conditions are passed on as input assignments to the consequent DRS.

In this format, DRS sequencing $\mathrm{K}_1 ; \mathrm{K}_2$ can easily be defined as relational composition:

DEFINITION 17. DRS sequencing

(i) $[\![\mathrm{K}_1 ; \mathrm{K}_2]\!] := \{\langle i, o\rangle | \exists m(\langle i, m\rangle \in \mathrm{K}_1 \wedge \langle m, o\rangle \in \mathrm{K}_2)\}$[15]

The relationship between DRT and models of computation has also been explored extensively within the framework of constructive/intuitionistic type theory [Martin-Löf, 1984]. We refer the reader to [Ahn and Kolb, 1990; Ranta, 1995; Fernando, 2001b; Fernando, 2001a]

3.2 Intensional Semantics, Propositions, Information States and Context Change Potential

Traditionally, the aim of model theoretic semantics has been to explicate meaning in terms of conditions of truth and reference. Often this goal is implemented via a two-step procedure: expressions of the object language (e.g. some fragment of English) are assigned a logical form or "semantic representation" — an expression belonging to some formal language. The model theoretic definition of truth conditions is then given directly for these semantic representations or logical forms. The truth conditions of an expression of the object language are in that case the truth conditions of the

[15]It is important to distinguish the DRS sequencing operation ";" from that of the merge of two or more DRSs. The merge $\mathrm{K}_1 \uplus \mathrm{K}_2$ of two DRSs is the DRS $\langle\ \mathrm{U}_{\mathrm{K}_1} \cup \mathrm{U}_{\mathrm{K}_2}$, $\mathrm{Con}_{\mathrm{K}_1} \cup \mathrm{Con}_{\mathrm{K}_2}\ \rangle$. Similarly, if $\mathcal{K}$ is a set of DRSs, then $\uplus\, \mathcal{K} = \langle \cup \{\ \mathrm{U}_{\mathrm{K}} — \mathrm{K} \in \mathcal{K}\ \}$, $\cup \{\ \mathrm{Con}_{\mathrm{K}} — \mathrm{K} \in \mathcal{K}\ \}\ \rangle$. Merge, unlike DRS sequencing, is a symmetric operation which obliterates any order between or among its arguments. It is an operation which is often useful in DRT, but it is alien to the dynamic perspective of QDL.

formal expressions assigned to it. This two-step procedure is reminiscent of DRT where we also assign formal representations (viz. DRSs) to bits of natural language and then state the truth definition as applying to DRSs. DRSs are assigned truth conditions, and the truth conditions of a DRS are to be understood as the truth conditions, and thus as the propositional content, of the bit of language it represents. But DRSs do more: they not only represent propositional content, but also provide the context against which new sentences in a discourse are interpreted. In DRT every new sentence in a discourse contributes to and in turn is interpreted against a continually evolving context. This new conception of meaning as context update and interpretation in context is the hallmark of "dynamic" semantics, which DRT and other early dynamic semantic theories such as File Change Semantics [Heim, 1982] initiated. One aspect of the contextual dependence of sentences in a cohesive text or dialogue is that in the bottom-up processing architecture the DRS constructed from a sentence which comes somewhere in the middle of a text will often be *improper*: it will contain occurrences of discourse referents which are free in the DRS itself (but belong to the universe of the context DRS; this happens whenever an anaphoric pronoun gets resolved in context, cf. examples (35), (39) and (43)). In these cases, it is only the merge of the new DRS with the context DRS to which the verification definition 10 and the truth definition 11 assign well defined truth conditions. The question that naturally arises at this point is: can we explicate the way in which the new sentence updates the context in which it is interpreted, in model theoretic terms, viz. by assigning it a function which maps the truth conditions of the context DRS to those of its update? When we move from an extensional model theory, of the kind we have assumed up to now, to an intensional one, in which it is possible to assign to every (proper) DRS the proposition (set of worlds) it expresses, then we can rephrase the above question as follows: can we associate with each improper DRS K a function CCP from propositions to propositions such that, if P_c is the proposition expressed by a context DRS K_c, then $\mathrm{CCP}_{\mathrm{K}}(\mathrm{P}_c)$ is the proposition expressed by the updated context, obtained through merging K_c with K? The answer to this question is negative. But it is nevertheless possible to achieve something that comes reasonably close to a positive answer: we can 'refine' the notion of the proposition expressed by a proper DRS K_c to that of the *information state* described by K_c, and can then assign to improper DRSs K update functions $\mathrm{CCP}_{\mathrm{K}}$ from information states to information states [Heim, 1982], such that if $\mathcal{I}_c$ is the information state described by the context DRS K_c and $\mathrm{CCP}_{\mathrm{K}}$ is the update function determined by K, then $\mathrm{CCP}_{\mathrm{K}}(\mathcal{I}_c)$ is the information state of the merge of K_c and K.

Below, we first present a simple, intensional semantics for the DRS language L defined in 2. We define the proposition expressed by a DRS K relative to $\mathcal{M}$ as the set of all possible worlds in $\mathcal{M}$ where K is true. We show that a simple version of CCP based on propositions is too coarse-grained to capture anaphoric dependencies (42), introduce the richer notion of information states and present a version of the CCP based on these.

To avoid certain notorious difficulties with existence and the denotation of names, we base the intensional model theory for the simple DRS language in 2 on models where all worlds come with the same universe (set of individuals) and where names denote once and for all (each name N denotes the same individual in every world of the model). Relations, however, are interpreted relative to particular worlds. We further assume that the accessibility relation between possible worlds is the universal relation (i.e. each world is accessible to itself and to each other world). An intensional model $\mathcal{M}$ is then defined as a triple $\langle \mathrm{W}_{\mathcal{M}}, \mathrm{U}_{\mathcal{M}}, \Im_{\mathcal{M}} \rangle$ as follows:

DEFINITION 18. An *intensional model* $\mathcal{M}$ is given by $\langle \mathrm{W}_{\mathcal{M}}, \mathrm{U}_{\mathcal{M}}, \Im_{\mathcal{M}} \rangle$, where:

(i) $\mathrm{W}_{\mathcal{M}}$ is a set of possible worlds

(ii) $\mathrm{U}_{\mathcal{M}}$ is a non-empty set

(iii) – for names, $\Im_{\mathcal{M}} : \mathrm{Name} \to \{\{d\} | d \in \mathrm{U}_{\mathcal{M}}\}$
– for n-ary relations, $\Im_{\mathcal{M}} : \mathrm{Rel}^n \to (\mathrm{W}_{\mathcal{M}} \to \mathcal{P}(\mathrm{U}^n))$

Verifying embeddings are defined globally, i.e. for some $\mathrm{X} \subseteq \mathrm{Ref}$, a verifying embedding g is defined as $g : \mathrm{X} \to \mathrm{U}_{\mathcal{M}}$ (and this assignment is understood as holding for all worlds, cf. clauses (ii)-(iv) of Defn. 19). An intensional semantics for DRSs and DRS conditions of L can now be defined as follows:

DEFINITION 19. Intensional semantics for DRSs and DRS conditions of L:

(i) $\langle g, h \rangle \models_{\mathcal{M},w} \langle \mathrm{U}, \mathrm{Con} \rangle$ iff $g \subseteq_{\mathrm{U}} h$ and for all $\gamma \in \mathrm{Con}$: $h \models_{\mathcal{M},w} \gamma$

(ii) $g \models_{\mathcal{M},w} \mathrm{x}_i = \mathrm{x}_j$ iff $g(\mathrm{x}_i) = g(\mathrm{x}_j)$

(iii) $g \models_{\mathcal{M},w} \mathrm{N}(\mathrm{x})$ iff $\{g(\mathrm{x})\} = \Im(\mathrm{N})$

(iv) $g \models_{\mathcal{M},w} \mathrm{P}(\mathrm{x}_1,..,\mathrm{x}_n)$ iff $\langle g(\mathrm{x}_1), .., g(\mathrm{x}_n) \rangle \in \Im(\mathrm{P})(w)$

(v) $g \models_{\mathcal{M},w} \neg\, \mathrm{K}$ iff there does not exist an h such that $\langle g, h \rangle \models_{\mathcal{M},w} \mathrm{K}$

(vi) $g \models_{\mathcal{M},w} \mathrm{K}_1 \vee \mathrm{K}_2$ iff there is some h such that $\langle g, h \rangle \models_{\mathcal{M},w} \mathrm{K}_1$ or there is some h such that $\langle g, h \rangle \models_{\mathcal{M},w} \mathrm{K}_2$

(vii) $g \models_{\mathcal{M},w} \mathrm{K}_1 \Rightarrow \mathrm{K}_2$ iff for all m such that $\langle g, m \rangle \models_{\mathcal{M}} \mathrm{K}_1$ there exists a k such that $\langle m, k \rangle \models_{\mathcal{M},w} \mathrm{K}_2$

A proper DRS K is true in $\mathcal{M}$ at a world w ($\models_{\mathcal{M},w}$ K) iff there exists an embedding h of U_K such that $\langle \Lambda, h\rangle \models_{\mathcal{M},w}$ K. Given a model $\mathcal{M}$, the proposition $[\![\mathrm{K}]\!]^p_{\mathcal{M}}$ expressed by a DRS K can now be defined as the set of all possible worlds in $\mathcal{M}$ in which K is true.

DEFINITION 20. Given a proper DRS K, the proposition $[\![\mathrm{K}]\!]^p_{\mathcal{M}}$ expressed by K relative to $\mathcal{M}$ is defined as:

$[\![\mathrm{K}]\!]^p_{\mathcal{M}} := \{w | \models_{\mathcal{M},w} \mathrm{K}\}$

The intensional semantics for DRT makes it possible to extend the repertoire of complex DRS conditions with intensional conditions whose verification at a world w may depend on the verification of the constituent DRSs at worlds other than w. Simple examples are conditions of the form $\Box$ K ("it is necessary that K") and $\Diamond$ K ("it is possible that K"), where K is a DRS. This extends L to a modal DRS language $\mathrm{L}_\Box$. Within the present intensional semantics we can state verification conditions for such DRS conditions which reflect Leibnitz' principle that necessary truth is truth in all possible worlds (while possible truth is truth in at least one possible world):

DEFINITION 21. *Verification of modal DRS conditions* of $\mathrm{L}_\Box$ in $\mathcal{M}$:

(i) $g \models_{\mathcal{M},w} \Box$ K iff *for all* $w' \in \mathrm{W}_{\mathcal{M}}$, $g \models_{\mathcal{M},w'}$ K

(ii) $g \models_{\mathcal{M},w} \Diamond$ K iff *for some* $w' \in \mathrm{W}_{\mathcal{M}}$, $g \models_{\mathcal{M},w'}$ K

Intensional models can also be used to formulate a semantics for DRSs that is dynamic in a somewhat different (and some say: stronger) sense than the versions given above. In DRT the construction of a semantic representation takes the form of an evolving context DRS where new pieces of discourse are interpreted against the available context and in turn update this context to a new context for the further following pieces to come. Given an already constructed context DRS $\mathrm{K}_{1,\ldots,n}$ for the first n sentences in a discourse, it is attractive to conceive of the dynamic semantic value of a DRS K_{n+1} for the next sentence S_{n+1} as transforming the semantic value $[\![\mathrm{K}_{1,\ldots,n}]\!]$ of the current context DRS into the new semantic value $[\![\mathrm{K}_{1,\ldots,n+1}]\!]$ for the extended context DRS $\mathrm{K}_{1,\ldots,n+1}$ which includes the information contributed by S_{n+1}. On this view, the semantic value of K_{n+1} would be a function from $[\![\mathrm{K}_{1,\ldots,n}]\!]$ to $[\![\mathrm{K}_{1,\ldots,n+1}]\!]$. The question is: what should these semantic values be? DRSs are associated with truth conditions and, given an intensional model, these, in their turn, define the proposition expressed by a DRS as the set of worlds where the DRS is true. Thus it might be tempting to build a dynamic semantics for DRT by defining the dynamic value of a DRS K_{n+1} as an operator which transforms the proposition expressed by the old context $\mathrm{K}_{1\ldots n}$ into the proposition expressed by the new context $\mathrm{K}_{1\ldots n+1}$. Formally this will give us, for each DRS K and model $\mathcal{M}$, a set of pairs of propositions relative to $\mathcal{M}$: where $\mathrm{Prop}_{\mathcal{M}} = \mathcal{P}(\mathrm{W}_{\mathcal{M}})$, $[\![\mathrm{K}]\!]^d_{\mathcal{M}} \subseteq$

$\text{Prop}_{\mathcal{M}} \times \text{Prop}_{\mathcal{M}}$. Note that such operators can only *add* information: for all K and $\text{P} \in \text{Prop}_{\mathcal{M}}$, $[\![\text{K}]\!]^d_{\mathcal{M}}(\text{P}) \subseteq \text{P}$.

This view may seem attractive as it attempts to explicate dynamic semantic values $[\![\cdot]\!]^d$ in terms of standard static semantic values $[\![\cdot]\!]^p$. There are, however, many examples (cf. (2) and (3)) that show that truth conditions alone (and dynamic semantic values based on functions from propositions to propositions) are insufficient to capture the dynamic meaning of semantic representations. Here we present a variant of a famous example due to Barbara Partee to illustrate the point

(42) (i). Exactly nine of the ten coins are in the bag and exactly one of the ten marbles is not. It is under the sofa.

(ii). Exactly nine of the ten marbles are in the bag and exactly one of the ten coins is not. It is under the sofa.

The DRSs for the first sentences in (42)(i) and (ii) are truth conditionally equivalent, i.e. they determine the same proposition. However, interpretation of the second sentence of (42.i,ii) in the context provided by the first sentence of (i) yields different results compared to its interpretation in the context provided by the first sentence of (ii). In (i) "It" refers to the missing marble, in (ii) to the missing coin.[16] Intuitively (as in examples (2) and (3)) the crucial difference between the first sentences of (42)(i) and (ii) above is not one of truth conditions but concerns which antecedents are made available for anaphoric reference in the following sentence.

In order to capture this difference, we need a more fine-grained notion of context than truth conditions and propositions provide. For the simple DRS fragment introduced up to now, the notion of an Information State [Heim, 1982; Groenendijk and Stokhof, 1990] provides the required granularity. For a proper DRS K and an intensional model $\mathcal{M}$, the information state $[\![\text{K}]\!]^s_{\mathcal{M}}$ records not just the worlds $w \in \text{W}_{\mathcal{M}}$ where K is true, but also the verifying embeddings f that make K true in w:

DEFINITION 22. Given a proper DRS K, the *information state* $[\![\text{K}]\!]^s_{\mathcal{M}}$ expressed by K relative to an intensional model $\mathcal{M}$ is defined as:

$$[\![\text{K}]\!]^s_{\mathcal{M}} := \{\langle w, f\rangle | \langle \Lambda, f\rangle \models_{\mathcal{M},w} \text{K}\}$$

Intuitively, verifying embeddings f for a given context DRS K record which discourse referents are available in the universe U_{K} as antecedents for anaphoric expressions occurring in sentences that are interpreted in the context of K. The embedding functions f occurring in the information state $\mathcal{I} = [\![\text{K}]\!]^s_{\mathcal{M}}$ expressed by a DRS K in $\mathcal{M}$ will all have the same domain,

[16] Partee's original example was:

(i) Exactly one of the ten marbles is not in the bag. It is under the sofa.

(ii) Exactly nine of the ten marbles are in the bag. It is under the sofa.

Here *it* is interpretable as referring to the missing marble in (i) but not in (ii).

viz. $\mathrm{U_K}$: if $\langle w, f\rangle \in \mathcal{I}$, then $\mathrm{Dom}(f) = \mathrm{U_K}$. We adopt this as a general constraint on information states (irrespective of whether they are the denotation of some DRS): for each information state $\mathcal{I}$ there is a set X of discourse referents such that for all $\langle w, f\rangle \in \mathcal{I}$, $\mathrm{Dom}(f) = \mathrm{X}$. X is called *the base of* $\mathcal{I}$ and denoted as $\mathrm{X}_\mathcal{I}$. Given a DRS K, the proposition $[\![\mathrm{K}]\!]^p_\mathcal{M}$ defined by K (i.e. the set of possible worlds in which K is true) can be recovered from the information state $[\![\mathrm{K}]\!]^s_\mathcal{M}$: $[\![\mathrm{K}]\!]^p_\mathcal{M} = \{w | \exists f \langle w, f\rangle \in [\![\mathrm{K}]\!]^s_\mathcal{M}\}$. It is clear that the mapping from information states to propositions is many to one: two sentences (such as the first sentences in (42.i) and (42.ii)) can express the same proposition but two different information states. Unlike propositions, information states record which discourse referents are provided by a context as potential antecedents for anaphoric NPs from sentences interpreted in this context.

DEFINITION 23. Given an intensional model $\mathcal{M}$, a DRS K and a set of discourse referents X we define

(i) $\mathcal{I}$ is an *information state relative to* $\mathcal{M}$ *and* X iff $\mathcal{I} \subseteq \{\langle w, f\rangle | \mathrm{Dom}(f) = \mathrm{X} \wedge \mathrm{Ran}(f) \subseteq \mathrm{U}_\mathcal{M} \wedge w \in \mathrm{W}_\mathcal{M}\}$

(ii) $\mathcal{I}$ is an *information state relative to* $\mathcal{M}$ iff there is an X such that $\mathcal{I}$ is an information state relative to $\mathcal{M}$ and X

(iii) when $\mathcal{I}$ is an *information state relative to* $\mathcal{M}$ *and* X, X is called the *base* of $\mathcal{I}$, and will sometimes be denoted as $\mathrm{X}_\mathcal{I}$

(iv) the empty information state $\Lambda^\mathcal{I}_\mathcal{M}$ relative to $\mathcal{M}$, $\Lambda^\mathcal{I}_\mathcal{M} := \{\langle w, \emptyset\rangle | w \in \mathrm{W}_\mathcal{M}\}$

(v) the *proposition* $\mathrm{Prop}(\mathcal{I})$ *determined by* $\mathcal{I}$: $\mathrm{Prop}(\mathcal{I}) := \{w | \exists f \langle w, f\rangle \in \mathcal{I}\}$

Given a context DRS K_i and a DRS K for a sentence interpreted in the context represented by K_i resulting in a new context K_o, the dynamic semantic value (i.e. the CCP) associated with K should transform the input context K_i to the output context K_o which results from updating K_i with K. K need not be a proper DRS as illustrated in the following example:

John owns a donkey. It loves him.

$$(43)\quad \begin{array}{|c|}\hline \text{x y} \\ \text{John(x)} \\ \text{donkey(y)} \\ \text{own(x,y)} \\ \hline \end{array} \uplus \begin{array}{|c|}\hline \text{z u} \\ \text{love(u,z)} \\ \text{z = x} \\ \text{u = y} \\ \hline \end{array} = \begin{array}{|c|}\hline \text{x y z u} \\ \text{John(x)} \\ \text{donkey(y)} \\ \text{own(x,y)} \\ \text{love(u,z)} \\ \text{z = x} \\ \text{u = y} \\ \hline \end{array}$$

$$\mathrm{K}_i \qquad\qquad \mathrm{K} \qquad\qquad \mathrm{K}_o$$

The context DRS K_i is a proper DRS but the DRS K is not since it contains occurrences of x and y free in K. K is anaphorically resolved in that equations z = x and u = y record with which antecedent discourse referents provided by the context DRS K_i the discourse referents z, u introduced into K by the anaphoric pronouns *it* and *him* are identified. But, as a consequence, K is not proper.

In the present case K can serve as an update of the context DRS K_i because the merge $\mathrm{K}_i \uplus \mathrm{K}$ of K and K_i is proper; or, put differently, because $\mathrm{FV(K)} \subseteq \mathrm{U}_{\mathrm{K}_i}$ ($= \mathrm{X}_{[\![\mathrm{K}_i]\!]^s_{\mathcal{M}}}$ for any model $\mathcal{M}$). This last condition is the key to the general principle underlying the characterization of the CCP of a DRS K in relation to a model $\mathcal{M}$: this should be a function that is defined on those information states $\mathcal{I}$ relative to $\mathcal{M}$ such that $\mathrm{FV(K)} \subseteq \mathrm{X}_{\mathcal{I}}$, and which in particular assigns to each such $\mathcal{I}$ which is of the form $[\![\mathrm{K}_i]\!]^s_{\mathcal{M}}$ the information state expressed by $\mathrm{K}_i \uplus \mathrm{K}$ as value. Generalising to arbitrary information states (i.e. abstracting away from the condition that $\mathcal{I}$ is expressed by some context DRS K_i) we get the following definition:

DEFINITION 24. The *Context Change Potential* (or the dynamic semantic interpretation) $[\![\mathrm{K}]\!]^d_{\mathcal{M}}$ of a DRS K relative to a model $\mathcal{M}$ is defined as a partial function from information states to information states such that :

(i) $[\![\mathrm{K}]\!]^d_{\mathcal{M}}$ is defined for those information states $\mathcal{I}$ relative to $\mathcal{M}$ such that $\mathrm{FV(K)} \subseteq \mathrm{X}_{\mathcal{I}}$

(ii) if $\mathcal{I}_i \in \mathrm{Dom}([\![\mathrm{K}]\!]^d_{\mathcal{M}})$, then $[\![\mathrm{K}]\!]^d_{\mathcal{M}}(\mathcal{I}_i) = \{\langle w, g\rangle | \exists f(\langle w, f\rangle \in \mathcal{I}_i \wedge \langle f, g\rangle \models_{\mathcal{M},w} \mathrm{K})\}$

For the example in (43) it is easy to see that $[\![\mathrm{K}]\!]^d_{\mathcal{M}}([\![\mathrm{K}_i]\!]^s_{\mathcal{M}}) = [\![\mathrm{K}_o]\!]^s_{\mathcal{M}}$, i.e. applying the dynamic semantic value associated with K to the information state expressed by K_i for the first sentence yields the information state expressed by K_o, the DRS representing the two sentences of (43) together.

Note that in case K is a proper DRS, $[\![\mathrm{K}]\!]^d_{\mathcal{M}}$ is a total function; put differently, if K is proper, then $[\![\mathrm{K}]\!]^d_{\mathcal{M}}$ is defined even for the empty information state $\Lambda^{\mathcal{I}}_{\mathcal{M}}$.

Useful Notions Relating Information States and CCPs

Information states can be ordered along two different dimensions. Intuitively, given two information states $\mathcal{I}$ and $\mathcal{I}'$ relative to the same base X, $\mathcal{I}'$ is at least as informative as $\mathcal{I}$ if $\mathcal{I}' \subseteq \mathcal{I}$. On the other hand, it is possible for an information state $\mathcal{I}'$ to be more informative than an information state $\mathcal{I}$, even though $\mathrm{Prop}(\mathcal{I}') = \mathrm{Prop}(\mathcal{I})$. For it may be that $\mathcal{I}'$ makes more discourse referents available than $\mathcal{I}$, i.e. $\mathrm{X}_{\mathcal{I}} \subset \mathrm{X}_{\mathcal{I}'}$ and moreover that whenever $\langle w, g\rangle \in \mathcal{I}'$, then there is an $f \subseteq g$ such that $\langle w, f\rangle \in \mathcal{I}$. This last condition can be used in a general definition of the relation "carries at least as much information as", which also applies to cases where $\mathrm{Prop}(\mathcal{I}) \neq \mathrm{Prop}(\mathcal{I}')$:

DEFINITION 25. Given two information states $\mathcal{I}$ and $\mathcal{I}'$, $\mathcal{I}'$ *carries at least as much information as* $\mathcal{I}$, in symbols $\mathcal{I} \preceq \mathcal{I}'$, iff $\forall w \forall g(\langle w, g\rangle \in \mathcal{I}' \rightarrow \exists f(\langle w, f\rangle \in \mathcal{I} \wedge f \subseteq g))$

Information states can be merged. We make use of this operation in Section 5 below.

DEFINITION 26. Let $\mathcal{M}$ be an intensional model and $\mathcal{S}$ a set of information states relative to $\mathcal{M}$. The *consistent merge of the* $\mathcal{I} \in \mathcal{S}$, denoted $\underline{\cup}\mathcal{S}$, is the information state defined by:

> $\underline{\cup}\mathcal{S} := \{\langle w, h\rangle|$ there exists a function F such that $\mathrm{Dom(F)} = \mathcal{S}$, for all $\mathcal{I} \in \mathcal{S}, \langle w, F(\mathcal{I})\rangle \in \mathcal{I}$ and $h = \cup\{F(\mathcal{I})|\mathcal{I} \in \mathcal{S}\}$ is a function. $\}$

N.B. Note that if $\underline{\cup}\mathcal{S} \neq \emptyset$, then the base of $\underline{\cup}\mathcal{S}$ is the union of the bases of the information states in $\mathcal{S}$, i.e. $\mathrm{X}_{\underline{\cup}\mathcal{S}} = \bigcup_{\mathcal{I}\in\mathcal{S}} \mathrm{X}_{\mathcal{I}}$.

When $\mathcal{S} = \{\mathcal{I}, \mathcal{I}'\}$ we also write $\mathcal{I}\underline{\cup}\mathcal{I}'$ in lieu of $\underline{\cup}\mathcal{S}$. Of particular importance are applications of consistent merge in cases where the bases of the members of $\mathcal{S}$ are disjoint, e.g. if $\mathcal{S} = \{\mathcal{I}, \mathcal{I}'\}$ and $\mathrm{X}_{\mathcal{I}} \cap \mathrm{X}_{\mathcal{I}'} = \emptyset$. In such applications the requirement that h must be a function is redundant.

In general, a CCP $\mathcal{J}$ relative to a model $\mathcal{M}$ is a function defined on some subset of the set of all information states relative to $\mathcal{M}$, which returns an information state relative to $\mathcal{M}$ for each information state in the domain. The CCPs $[\![\mathrm{K}]\!]^d_{\mathcal{M}}$ determined by some DRS K fit this general description, but they satisfy further conditions:

(i) whether an information state $\mathcal{I}$ belongs to the domain of such a CCP $\mathcal{J}$ depends exclusively on its base $\mathrm{X}_{\mathcal{I}}$. More precisely, there is a set of discourse referents $\mathrm{J}_{\mathcal{J}}$ such that $\mathcal{I} \in \mathrm{Dom}(\mathcal{J})$ iff $\mathrm{J}_{\mathcal{J}} \subseteq \mathrm{X}_{\mathcal{I}}$. We call $\mathrm{J}_{\mathcal{J}}$ the *referential presupposition* of $\mathcal{J}$.

(ii) $\mathcal{J}$ has a *base* $\mathrm{X}_{\mathcal{J}}$, a set of discourse referents such that if $\mathcal{J}$ is defined for $\mathcal{I}$, then $\mathrm{X}_{\mathcal{J}(\mathcal{I})} = \mathrm{X}_{\mathcal{I}} \cup \mathrm{X}_{\mathcal{J}}$.

(iii) $\mathcal{J}$ is *distributive*, i.e. if $\mathcal{J}$ is defined for $\mathcal{I}$, then $\mathcal{J}(\mathcal{I}) = \cup\{\mathcal{J}(\{\langle w, f\rangle\})|\langle w, f\rangle \in \mathcal{I}\}$.

We call CCPs which satisfy conditions (i) to (iii) *regular* CCPs. These informal stipulations are summarised in in Definition 27:

DEFINITION 27. Let $\mathcal{M}$ be an intensional model, $\mathcal{J}$ a CCP relative to $\mathcal{M}$ and $\mathrm{X}_{\mathcal{J}}$ a set of discourse referents. $\mathcal{J}$ is *regular with base* $\mathrm{X}_{\mathcal{J}}$ iff

(i) for arbitrary information states $\mathcal{I}$ relative to $\mathcal{M}$, $\mathcal{I} \in \mathrm{Dom}(\mathcal{J})$ iff $\mathrm{X}_{\mathcal{J}} \subseteq \mathrm{X}_{\mathcal{I}}$

(ii) for $\mathcal{I} \in \mathrm{Dom}(\mathcal{J})$, $\mathcal{I} \preceq \mathcal{J}(\mathcal{I})$

Note that if $\mathcal{J}$ is both regular and total, then $\mathcal{J}$ is defined on all information states relative to $\mathcal{M}$:

DEFINITION 28. Total Context Change Potential

A Context Change Potential $\mathcal{J}$ is total iff $\mathcal{J}(\Lambda^{\mathcal{I}}_{\mathcal{M}})$ is defined.

The notion of the proposition expressed by a DRS K relative to a model $\mathcal{M}$ and that of the information state expressed by K have so far been defined exclusively for proper DRSs. But they can be readily generalised to improper DRSs by making them dependent on assignments to the free discourse referents of the DRS. For instance, when K is a DRS and g is a map from FV(K) into $\mathrm{U}_{\mathcal{M}}$, then the proposition expressed by K in $\mathcal{M}$ relative to g can be defined as the set of those worlds w of $\mathcal{M}$ for which there is an $g \subseteq_{\mathrm{U}_{\mathrm{K}}} h$ such that $\langle g, h\rangle \models_{\mathcal{M},w} \mathrm{K}$ (see Definition 19). The notion of an information state relative to $\mathcal{M}$ can be generalised analoguously. The formal charachterisations are given in the next definition.

DEFINITION 29. Let $\mathcal{M}$ be an intensional model, K a possibly improper DRS, g an assignment of FV(K) in $\mathcal{M}$. Then

(i) the *proposition expressed by K relative to* $\mathcal{M}$ *and* g, $[\![\mathrm{K}]\!]^p_{\mathcal{M},g}$, is defined by
$[\![\mathrm{K}]\!]^p_{\mathcal{M},g} = \langle w \in \mathrm{W}_{\mathcal{M}}|(\exists h)(g \subseteq_{\mathrm{U}_{\mathrm{K}}} h \wedge \langle g, h\rangle \models_{\mathcal{M},g} \mathrm{K}\rangle$

(ii) the information state expressed by K relative to $\mathcal{M}$ and g, $[\![\mathrm{K}]\!]^s_{\mathcal{M},g}$, is defined by
$[\![\mathrm{K}]\!]^s_{\mathcal{M},g} = \{\langle w, f\rangle|g \subseteq_{\mathrm{U}_{\mathrm{K}}} f \wedge \langle g, f\rangle \models_{\mathcal{M},w} \mathrm{K}\}$

For DRSs K from the extensional DRS-languages we have so far considered there is a close relation between $[\![\mathrm{K}]\!]^d_{\mathcal{M}}$ and $[\![\mathrm{K}]\!]^s_{\mathcal{M},g}$. Suppose that $\mathcal{I}$ is an information state in the domain of $[\![\mathrm{K}]\!]^d_{\mathcal{M}}$ and that $\langle w, g\rangle \in \mathcal{I}$. Then we have for any f such that $g \subseteq_{\mathrm{U}_{\mathrm{K}}} f$:

(44) $\langle w, f\rangle \in [\![\mathrm{K}]\!]^d_{\mathcal{M}}(\mathcal{I})$ iff $\langle w, f\rangle \in [\![\mathrm{K}]\!]^s_{\mathcal{M},g}$.

This property is closely connected with the fact that the context change potentials defined by such DRSs are extensional in the sense described below. For any two information states $\mathcal{I}$ and $\mathcal{I}'$ relative to $\mathcal{M}$ with the same base (i.e. $\mathrm{X}_{\mathcal{I}} = \mathrm{X}_{\mathcal{I}'}$) we say that $\mathcal{I}$ and $\mathcal{I}'$ coincide on $w \in \mathrm{W}_{\mathcal{M}}$ iff $\{f|\langle w, f\rangle \in \mathcal{I}\}$ $= \{f : \langle w, f\rangle \in \mathcal{I}'\}$.

A CCP $\mathcal{J}$ relative to $\mathcal{M}$ is called *extensional* iff whenever $w \in \mathrm{W}_{\mathcal{M}}$, $\mathcal{I}$, $\mathcal{I}' \in \mathrm{Dom}(\mathcal{J})$, $\mathrm{X}_{\mathcal{I}} = \mathrm{X}_{\mathcal{I}'}$ and $\mathcal{I}$ and $\mathcal{I}'$ coincide on w, then $\mathcal{J}(\mathcal{I})$ and $\mathcal{J}(\mathcal{I}')$ coincide on w. It is not hard to verify that when K is a DRS from the extensional DRS language defined above (which does not contain $\Box$ and $\Diamond$), then $[\![\mathrm{K}]\!]^d_{\mathcal{M}}$ is extensional.

For certain purposes it is convenient to be able to make use of λ-abstracts over DRSs. As in Intensional Higher Order Logic ([Gallin, 1975]) we admit two kinds of λ-abstraction.

(i) extensional λ-abstraction over free discourse referents in a DRS.

(ii) intensional abstraction denoted by the operator $^\wedge$, which is de facto an abstraction operator over worlds.

It has proved convenient to assume that λ-abstraction over discourse referents may involve any non-empty subset $\{\mathrm{x}_1 ,\ldots, \mathrm{x}_n\}$ of the free discourse referents of the DRS (rather than just a single discourse referent). The definitions follow the pattern familiar from the model theory for formalisms with abstraction operators, and as such they are unsurprising. The only complication we are facing is that we have defined several types of semantic values for the objects to which these operators apply, viz. DRSs. A similar variety of options does in principle exist for the terms which we get by applying an abstraction operator to a DRS. We limit our attention here to truth values, propositions and information states. The formal definitions are given in Definition 30.

DEFINITION 30. Let $\mathcal{M}$ be an intensional model, K a DRS and let $\mathrm{x}_1 ,\ldots, \mathrm{x}_n \in \mathrm{FV}(\mathrm{K})$.

(a) Let $w \in \mathrm{W}_{\mathcal{M}}$, g an assignment in $\mathcal{M}$ on $\mathrm{FV}(\mathrm{K}) \setminus \{\mathrm{x}_1 ,\ldots, \mathrm{x}_n\}$.

(i) $[\![\lambda\{\mathrm{x}_1, ..., \mathrm{x}_n\}.\mathrm{K}]\!]_{\mathcal{M},w,g}$ is that function from $(\mathrm{U}_{\mathcal{M}})^n$ to truth values which is given by:
if $\mathrm{a}_1 ,\ldots, \mathrm{a}_n \in \mathrm{U}_{\mathcal{M}}$, then $[\![\lambda\{\mathrm{x}_1, \ldots, \mathrm{x}_n\}.\mathrm{K}]\!]_{\mathcal{M},w,g}(\langle \mathrm{a}_1, \ldots, \mathrm{a}_n\rangle) =$ 1 iff $(\exists h)(g \cup \{\langle \mathrm{x}_1, \mathrm{a}_1\rangle, \ldots, \langle \mathrm{x}_n, \mathrm{a}_n\rangle\} \subseteq_{\mathrm{U}_\mathrm{K}} h \wedge \langle g, h\rangle \models_{\mathcal{M},w} \mathrm{K})$

(ii) $[\![\lambda\{\mathrm{x}_1, \ldots, \mathrm{x}_n\}.\mathrm{K}]\!]^p_{\mathcal{M},w,g}$ is that function from $(\mathrm{U}_{\mathcal{M}})^n$ to propositions relative to $\mathcal{M}$ such that for $\mathrm{a}_1 ,\ldots, \mathrm{a}_n \in \mathrm{U}_{\mathcal{M}}$:
$[\![\lambda\{\mathrm{x}_1, \ldots, \mathrm{x}_n\}.\mathrm{K}]\!]^p_{\mathcal{M},w,g}(\langle \mathrm{a}_1, \ldots, \mathrm{a}_n\rangle) =$

$\{w' \in \mathrm{W}_{\mathcal{M}} | (\exists h)(g \cup \{\langle \mathrm{x}_1, \mathrm{a}_1\rangle, \ldots, \langle \mathrm{x}_n, \mathrm{a}_n\rangle\} \subseteq_{\mathrm{U}_\mathrm{K}} h \wedge \langle g, h\rangle \models_{\mathcal{M},w'}$ K))}

(iii) $[\![\lambda\{\mathrm{x}_1, \ldots, \mathrm{x}_n\}.\mathrm{K}]\!]^s_{\mathcal{M},w,g}$ is that function from $(\mathrm{U}_{\mathcal{M}})^n$ to information states relative to $\mathcal{M}$ such that for $\mathrm{a}_1, \ldots, \mathrm{a}_n \in \mathrm{U}_{\mathcal{M}}$:
$[\![\lambda\{\mathrm{x}_1, \ldots, \mathrm{x}_n\}.\mathrm{K}]\!]^p_{\mathcal{M},w,g}(\langle \mathrm{a}_1, \ldots, \mathrm{a}_n\rangle) =$
$\{\langle w', f\rangle | w' \in \mathrm{W}_{\mathcal{M}} \wedge g \cup \{\langle \mathrm{x}_1, \mathrm{a}_1\rangle, ..., \langle \mathrm{x}_n, \mathrm{a}_n\rangle\} \subseteq_{\mathrm{U}_\mathrm{K}} f \wedge \langle g, f\rangle \models_{\mathcal{M},w'}$ K))}

(b) We consider two possible operands for the intensional abstraction operator $^\wedge$, (i) DRSs and (ii) λ-abstracts over DRSs. Moreover, we only define the effect of $^\wedge$ as a proposition forming operator, in the following sense: If the operand is a DRS K, we consider $^\wedge$ as forming a term denoting the proposition expressed by K (relative to some assignment, when K is improper). If the operand is a λ-abstract $\lambda\{\mathrm{x}_1, \ldots, \mathrm{x}_n\}.\mathrm{K}$ then the result of applying $^\wedge$ is a term which denotes a propositional function, i.e. a function which for each combination of objects $\mathrm{a}_1,...,\mathrm{a}_n \in \mathrm{U}_{\mathcal{M}}$ returns a proposition relative to $\mathcal{M}$ as value. Again the definitions are unsurprising.

(iv) $[\![^\wedge\mathrm{K}]\!]_{\mathcal{M},w,g} = \{w' | w' \in \mathrm{W}_{\mathcal{M}} \wedge (\exists h)(g \subseteq_{\mathrm{U}_\mathrm{K}} h \wedge \langle g, h\rangle \models_{\mathcal{M},w'} K)\}$

(v) $[\![^\wedge\lambda\{x_1, ..., x_n\}.\mathrm{K}]\!]_{\mathcal{M},w,g}$ is that function from $(\mathrm{U}_{\mathcal{M}})^n$ to propositions relative to $\mathcal{M}$ such that such that for $\mathrm{a}_1, \ldots, \mathrm{a}_n \in \mathrm{U}_{\mathcal{M}}$:
$[\![\lambda\{\mathrm{x}_1, ..., \mathrm{x}_n\}.\mathrm{K}]\!]^p_{\mathcal{M},w,g}(\langle \mathrm{a}_1, \ldots, \mathrm{a}_n\rangle) =$
$\{w' \in \mathcal{M} | (\exists h)(g \cup \{\langle \mathrm{x}_1, \mathrm{a}_1\rangle, ..., \langle \mathrm{x}_n, \mathrm{a}_n\rangle\} \subseteq_{\mathrm{U}_\mathrm{K}} h \wedge \langle g, h\rangle \models_{\mathcal{M},w'}$ K))}

It would be possible to generalise these definitions to a full fledged Higher Order Intensional Dynamic Logic. But the generalised definitions become fairly abstract, and we have not been able to envisage much use for them in relation to the aspects of DRT discussed in this survey.

Abstraction of either kind is also possible for DRS conditions. We can reduce such applications by identifying a DRS condition γ with the DRS $\langle \emptyset, \{\gamma\}\rangle\}$. In later parts of this survey (in particular in Section 3.5) we will need in particular terms of the form $^\wedge$x.γ, where x is a free variable of γ. These are short for $^\wedge\lambda\{\mathrm{x}\}.\gamma$, or, more precisely, $^\wedge\lambda\{\mathrm{x}\}.\langle \emptyset, \{\gamma\}\rangle$.

3.3 Generalised Quantifiers

One of the central tenets of DRT is that certain expressions which earlier theories treat as quantifiers should not be treated in this way. In particular, indefinites, DRT claims, should not be treated as quantificational expressions, but rather as terms, and thus in a manner that aligns them more closely with definite noun phrases than with the (genuinely) quantifying NPs. What quantificational force individual occurrences of indefinites may seem to have is, it is argued, an indirect effect — a kind of side effect due to the interactions with such operators as negation or implication.

Connected with this perspective is the fact that the orginal DRT formalism while expressively equivalent to first order predicate logic, nevertheless differs from it importantly in the way in which its "formulas" (i.e. the DRSs) parcel the information they contain. In particular, DRT differs from first order logic in that it doesn't make a principled distinction between sentential and quantificational operators. In fact, the original formalism didn't have any quantifiers as such. What comes nearest to a quantifier in this system is the implication operator $\Rightarrow$. But even this operator is not a quantifier strictly speaking. It acts like a quantifier only when at least one of the DRSs it connects has a non-empty universe, and what kind of quantification it represents then further depends on which of those universes is non-empty: As we saw in Section 3.1, the force of an implicational condition like that in (45) is that of a plain sentential conditional if $U_1 = U_2 = \emptyset$; of a restricted universal quantifier if $U_1 \neq \emptyset$ and $U_2 = \emptyset$; of a conditionalized existential quantifier if $U_1 = \emptyset$ and $U_2 \neq \emptyset$; and of a quantificational complex in which some universal quantifiers are followed by some existential quantifiers, if $U_1 \neq \emptyset$ and $U_2 \neq \emptyset$.

(45) $\langle U_1, \mathrm{Con}_1 \rangle \Rightarrow \langle U_2, \mathrm{Con}_2 \rangle$

There is arguably a sense in which $\Rightarrow$ is the universal quantifier of the original DRT formalism. For one thing it is used in the representation of the universal quantifiers that are part of the natural language fragment for which the first DRT accounts provided a systematic analysis, i.e. NPs with the determiner *every*. For instance, as discussed in Section 2, the universally quantified sentence

(46) Every farmer who uses a tractor has a neigbour with whom he shares it.

is represented as in (47).

(47)
$$\begin{array}{|c|}\hline \\ \begin{array}{|c|}\hline \mathrm{x\ y} \\ \mathrm{farmer(x)} \\ \mathrm{tractor(y)} \\ \mathrm{owns(x,y)} \\ \hline \end{array} \Rightarrow \begin{array}{|c|}\hline \mathrm{z\ u} \\ \mathrm{u = y} \\ \mathrm{neighbour(z,x)} \\ \mathrm{share(x,u,z)} \\ \hline \end{array} \\ \\ \hline \end{array}$$

But (47) shows that even in those cases where the universe U_1 is non-empty and $\Rightarrow$ consequently involves universal quantification of some sort, it does not quite behave like the universal quantifier of predicate logic in its standard form. Even if in addition $U_2 = \emptyset$ there still are the following two differences: (i) $\Rightarrow$ operates on two formula-like arguments (the DRSs to the left and right of the arrow) rather than one; (ii) $\Rightarrow$ is "unselective", binding all the discourse referents in the universe of the first argument DRS.

(i) is in keeping with a now well-neigh universally accepted view of how quantification in natural language typically works: it involves an operator which takes two predicates as arguments, the first called its *restrictor*, and the second its *(nuclear) scope*. In particular, when quantification is expressed by a noun phrase such as the subject of (46), it is the common noun phrase of this NP that acts as restrictor, while the scope of the quantifier is provided by the remainder of the clause in which the NP occurs as a constituent. Structures of this sort have been studied extensively within generalised quantifier theory (see [Westerstahl, 1989a]), in which quantifiers are analysed as *variable binding operators* — operators which bind one or more variables and whose arguments are formulas that, in the typical case, contain free occurrences of the variable or variables the operator binds. The special case of immediate interest is that represented in (48), of an operator Q which binds one variable and takes two formula-arguments.

(48) $\mathrm{Qx}(\mathrm{A}(\mathrm{x}), \mathrm{B}(\mathrm{x}))$

The standard interpretation of such a quantifier Q is as a relation R(Q) between sets: (48) is true if the set of x's satisfying A stands in the relation R(Q) to the set of x's satisfying B. In particular, the universal quantifier is interpreted as inclusion: if Q is the universal quantifier, then (48) is true iff the set of the A's is included in the set of the B's. We will see presently in what sense DRT's implication operator conforms to this analysis of universal quantification.

(ii) is more controversial. It was argued in [Lewis, 1975] that quantification in natural language is unselective: the quantificational operator binds whatever bindable variables turn up within its immediate scope; in principle there is no upper bound to the number of variables that a single operator can bind. Original DRT (and likewise File Change Semantics, see [Heim, 1982]) adopted the unselective analysis of quantification because of the attractive solution that it seems to offer to the "donkey problem" — how to account for the fact that in a sentence like (46) the indefinite *a tractor* inside the quantifying subject NP has the apparent force of a universal quantifier whose scope extends beyond the NP and includes all other sentence material (see Section 2, [Kamp, 1981a]).

The generalised quantifier semantics described for (48) can be naturally adapted to the case of the unselective universal "quantifier" $\Rightarrow$: a DRS condition governed by $\Rightarrow$ is true if a certain set associated with the first argument (i.e. the left DRS) is included in the corresponding set associated with the second (the right DRS). But in view of the unselectiveness of $\Rightarrow$ we need to adjust the definition of the sets. Instead of the set of objects satisfying the first argument of $\Rightarrow$ we must now consider the set of possible assignments of objects to all the discourse referents in the universe of the left DRS — in the case of (47) this is a set of pairs of objects $\langle a, b\rangle$, where a is assigned to the discourse referent x and b to y, and where this assignment satisfies the conditions on the left. Similarly, the second set should consist of those assignments that satisfy the first argument and which can be extended to an assignment which includes the discourse referents on the right and satisfies the second argument of $\Rightarrow$ — in the case of (47) these are the pairs $\langle a, b\rangle$ which satisfy the conditions on the left and can be extended to tuples $\langle a, b, c, d\rangle$ with c assigned to z and d to u, which satisfy the conditions on the right. It is easily seen that (47) is true according to Def. 10 iff the first of these sets is included in the second.

Duplex Conditions and the Proportion Problem

It was soon noted that unselectivity leads to problems with non-universal quantifiers. This is the so-called "proportion problem" [Kadmon, 1987, Chapter 10]. The problem is easiest to explain for the case of the quantifier *most*. It is quite generally held that a sentence like (49) is true if the number of farmers that are rich exceeds the number of farmers that are not rich. (More generally and formally: *Most A's are B's* is true iff the cardinality of the set $A \cap B$ is bigger than that of the set $A \setminus B$). But what are the truth conditions of sentence (50)?

(49) Most farmers are rich.

(50) Most farmers who use a tractor share it with a neighbour.

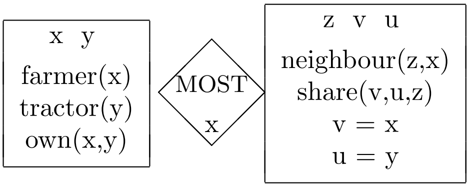
By analogy with what we have just said about (46) one would expect the following:

(51) (50) is true iff the number of assignments $\langle a, b\rangle$ to x, y which satisfy the conditions on the left of (47) and can be extended to an assignment $\langle a, b, c, d\rangle$ which satisfies the conditions on the right exceeds the number of assignments $\langle a, b\rangle$ which satisfy the conditions on the left but which cannot be so extended.

However, linguistic intuition tells us that this cannot be right; in a case where there are 19 farmers who each use just one tractor and share this tractor with some neighbour, while the 20-th farmer uses 25 tractors none of which he shares with anyone, (50) seems intuitively true (it is 19 against 1!), but the condition we have just stated predicts it to be false, as there are 19 pairs $\langle a, b\rangle$ of the first kind and 25 of the second.

The trouble with (51) is that it counts numbers of assignments (here pairs consisting of a farmer and a tractor) rather than just the numbers of farmers. What should be counted are just the satisfiers of the one variable which, in the standard generalised quantifier format (48), is bound by the quantifier. In order to correct (51) so that it conforms to this intuition, the discourse representation of (50) (and by parity those of (46) and other quantified sentences) should mark the "bound variable" in some way, so that it can be distinguished from the other discourse referents on the left. To this end DRT has adopted the so-called duplex conditions. An example is the duplex condition representing (50), given in (52).

(52)



In general, a duplex condition consists of three parts, (i) the restrictor DRS, (ii) the scope DRS and (iii) the quantifier part. (i) and (ii) are as in the earlier DRT representations of quantification (cf. the left hand side DRS and right hand side DRS in 47); the quantifier part consists of a quantifier (*most*, *every*, *many*, etc.) and a discourse referent (corresponding to the bound variable in (48)).

There is one aspect of the duplex notation which requires comment. This is the simultaneous occurrence of the discourse referent x as a constituent of the quantifier part and as a member of the universe of the restrictor DRS. This two-fold occurrence could create the impression that x is bound "twice over", something which would be logically incoherent. But this is not what is intended. Only the occurrence of x as constituent of the quantifier acts as a binding. In fact, we could eliminate the occurrence in the universe of the restrictor DRS provided we adjust the definition of accessibility (see 3.1) by stipulating that a discourse referent occurring as constituent of the quantifier part of a duplex condition acts, for the purpose of accessibility, as if it was a member of the universe of the restrictor DRS. The duplex notation exemplified in (52), in which the operator-bound discourse referent is added explicitly to the restrictor universe, obviates the need for this stipulation. The presence of this discourse referent within the quantifier part then sets it apart from the other members of the restrictor universe, in a way that is made explicit in the verification conditions for duplex conditions.

It is through its quantifier part that (52) provides us with the distinction which we need if we are to revise the truth condition (51) so that it conforms to intuition. But how should (51) be changed? This is not obvious. In fact, what exactly does (50) assert? Does it say that the majority of the farmers who use a tractor has the property that they share every tractor they use with some neighbor; or does the sentence require that there be a majority who share at least one of the tractors they use? There is surprisingly little agreement on this question, and the linguistic debate which sentences of the general form of (50) have, or prefer, the one interpretation and which the other, remains inconclusive to this day. (See [Chierchia, 1991], [Rooth, 2005].)

This is not the place to take sides in this debate. We only note that it has led to the names for the two readings which are now in general use; the first reading, according to which the second of the two sets consists of the farmers who share *all* the tractors they use is called the *strong* reading (of donkey sentences); the other, according to which the set consists of farmers who share some of the tractors they use, is the *weak* reading.

When turning (51) into a verification condition for duplex conditions like that in (52), we must see to it that the new condition accords with our intuitions about the proportion problem; but this still leaves us both the option for the strong reading and that for the weak reading. The second option, given in (54), is formally simpler and more elegant. It can be stated as follows. Recall that each duplex condition has a restrictor DRS K_r and a scope DRS K_s, and that its quantifier part binds one discourse referent, say x. For simplicity let us assume that $\mathrm{FV}(\mathrm{K}_r) = \emptyset$ and that $\mathrm{FV}(\mathrm{K}_s) \subseteq \mathrm{U}_{\mathrm{K}_r}$. Define:

(53) i. $S_r :=$ the set of all objects a such that there is an assignment h of objects to the discourse referents in the universe of K_r which assigns a to x and verifies the conditions of K_r;

ii. $S_s :=$ the set of all a such that there is an assignment h to the discourse referents of the universes of K_r and K_s which assigns a to x and verifies the conditions of K_r and K_s.

(54) (52) is true iff $|S_s| > |S_r \setminus S_s|$

The corresponding condition for the strong reading can be stated in a similar form:

(55) (52) is true iff $|S'_s| > |S_r \setminus S'_s|$

Superficially, this looks much like (54), but the definition of S'_s is more complex and awkward than that of S_s:

(56) iii. $S'_s :=$ the set of all a such that (i) there is an assignment h to the discourse referents of the universe of K_r which assigns a to x and verifies the conditions of K_r and (ii) for every assignment h to the discourse referents of the universe of K_r which assigns a to x and verifies the conditions of K_r there is an assignment k to the discourse referents of the universe K_s which extends h and verifies the conditions of K_s.

Some linguists have taken the complexity of (56.iii) as an indication that the strong reading cannot be the primary interpretation of donkey sentences such as (50). (See e.g. [Chierchia, 1993]).

N.B. (54) and (55) come close to what is needed when duplex conditions are added to the DRS language defined in Section 3.1. But the additional truth clauses for such connectives as supplements of Def. 10 require a slightly more complicated form. For instance, for the weak reading of (52) the new clause is

(57) $$g \models_{\mathcal{M}} \mathrm{K}_r \left\langle \begin{matrix}\text{MOST}\\ \text{x}\end{matrix} \right\rangle \mathrm{K}_s \text{ iff } |S_s^g| > |S_r^g \setminus S_s^g|$$

where S_r^g = the set of $a \in \mathrm{U}_{\mathcal{M}}$ sucht that there is an assignment h such that $g \subseteq_{\mathrm{U}_{\mathrm{K}_r}} h$, $h(\mathrm{x}) = a$ and h verifies the conditions of K_r.

And analoguously for S_s^g (see (53.ii).)

The truth conditions for the strong reading of duplex conditions with the operator MOST is obtained from (55) and (56) in the same way that (56) can be obtained from (53) and (54). Duplex conditions for other quantifiers will also give rise to truth conditions for either the weak or the strong reading.

Natural languages have various constructions for expressing quantification. No less prominent than quantifying noun phrases are adverbs of quantification — *always*, *never*, *often*, *mostly* and so on. In fact it was quantificational adverbs which led Lewis [1975] to his proposal of non-selective quantification; and as an analysis of adverbial quantification this proposal stands up much better than it does for quantification by means of noun phrases; adverbial quantification is much less vulnerable to objections connected with the proportion problem.

For instance, consider the adverbial analogue (58) of (50):

(58) Mostly when a farmer uses a tractor, he shares it with a neighbour.

In the scenario we considered in connection with (50), (50) itself seemed unequivocally true. But for (58) this is much less evident. Here a good case can be made for the claim that it is the numbers of farmer-tractor pairs which are to be counted, and not just the farmers.

This judgement seems to show that adverbial quantification can involve binding of several variables by the same quantifier. There has been discussion in the literature whether even these cases should be analysed as involving a single bound variable, ranging over "occasions", or "situations", where such occasions or situations may have several participants. (E.g. (58) would be analysed as quantifying over situations which each involve a farmer and a tractor that farmer uses.)[17] In the light of the commitment we have already made to duplex conditions, in which the left hand side DRS universe may contain discourse referents besides the one which is bound by the quantifier, this debate loses much of its urgency and we will assume without further argument that quantificational binding of more than one variable is indeed possible. Thus for (58) we assume a representation of the form given in (59).

(59)
$$\boxed{\boxed{\begin{array}{c}\mathrm{x\ y}\\ \hline \mathrm{farmer(x)}\\ \mathrm{tractor(y)}\\ \mathrm{uses(x,y)}\end{array}} \;\Big\langle\begin{array}{c}\mathrm{MOST}\\ \mathrm{x,y}\end{array}\Big\rangle\; \boxed{\begin{array}{c}\mathrm{z\ u}\\ \hline \mathrm{u = y}\\ \mathrm{neighbour(z,x)}\\ \mathrm{share(x,u,z)}\end{array}}}$$

More generally, we assume that duplex conditions representing sentences with the quantifier *most(ly)* instantiate the following schematic form.

(60)
$$\boxed{\boxed{\begin{array}{c}\mathrm{x}_1 \ldots \mathrm{x}_n\ \mathrm{y}_1 \ldots \mathrm{y}_m\\ \hline \mathrm{C}_1(\mathrm{x}_1,\ldots,\mathrm{x}_n,\mathrm{y}_1,\ldots,\mathrm{y}_m)\\ \vdots\\ \mathrm{C}_r(\mathrm{x}_1,\ldots, \mathrm{x}_n,\mathrm{y}_1,\ldots, \mathrm{y}_m)\end{array}} \;\Big\langle\begin{array}{c}\mathrm{MOST}\\ \mathrm{x}_1\ldots\mathrm{x}_n\end{array}\Big\rangle\; \boxed{\begin{array}{c}\mathrm{z}_1 \ldots \mathrm{z}_k\\ \hline \mathrm{D}_1(\mathrm{x}_1,\ldots,\mathrm{x}_n,\mathrm{y}_1,\ldots,\mathrm{y}_m,\mathrm{z}_1,\ldots,\mathrm{z}_k)\\ \vdots\\ \mathrm{D}_s(\mathrm{x}_1,\ldots,\mathrm{x}_n,\mathrm{y}_1,\ldots,\mathrm{y}_m,\mathrm{z}_1,\ldots,\mathrm{z}_k)\end{array}}}$$

A duplex condition of the general form of (60) is verified by an embedding f iff
$\mid S_2 \mid$ ¿ $\mid S_1 \setminus S_2 \mid$, where S_1, S_2 are as defined as follows:

$$S_1 = \{\langle \mathrm{a}_1,\ldots,\mathrm{a}_n\rangle \text{ — } (\exists g)(f \subseteq_{\mathrm{U}} g \wedge \textstyle\bigwedge_{i=1}^{n} g(\mathrm{x}_i) = \mathrm{a}_i \wedge \bigwedge_{j=1}^{r} g \models_{\mathcal{M}} \mathrm{C}_j(\mathrm{x}_1,\ldots,\mathrm{x}_n,\mathrm{y}_1,\ldots,\mathrm{y}_m))\}$$
where $\mathrm{U} = \{\mathrm{x}_1,\ldots,\mathrm{x}_n,\mathrm{y}_1,\ldots,\mathrm{y}_m\}$

$$S_2 = \{\langle \mathrm{a}_1,\ldots,\mathrm{a}_n\rangle \text{— } (\exists g)(f \subseteq_{\mathrm{U}} g \wedge \textstyle\bigwedge_{i=1}^{n} g(\mathrm{x}_i) = \mathrm{a}_i \wedge \bigwedge_{j=1}^{r} g \models_{\mathcal{M}} \mathrm{C}_j(\mathrm{x}_1,\ldots,\mathrm{x}_n,\mathrm{y}_1,\ldots,\mathrm{y}_m) \wedge (\exists h)(g \subseteq_{\mathrm{V}} h \wedge \bigwedge_{t=1}^{s} h \models_{\mathcal{M}} \mathrm{D}_t(\mathrm{x}_1,\ldots, \mathrm{x}_n,\mathrm{y}_1,\ldots,\mathrm{y}_m, \mathrm{z}_1,\ldots,\mathrm{z}_k)))\}$$
where $\mathrm{V} = \{\mathrm{z}_1,\ldots,\mathrm{z}_k\}$.

Even more generally, the duplex conditions may have some other operator Q occupying the position of MOST in (60). The truth conditions of such

[17]In Section 3.5 we will consider frequency adverbs like *mostly* more closely, albeit only in their role as quantifiers over times. We have just seen that such adverbs have other uses as well — (58) is a case in point, as it need not be interpreted as a case of temporal quantification, and on its most natural interpretation it is not. Nevertheless, the analysis we will consider there of the temporal uses of such adverbs is instructive from the point at issue here. For it shows how the times t over which the quantifier ranges can serve as representations for groups individuals — those individuals which stand in a certain relationship at t.

duplex conditons will be given by some relation between the sets S_1 and S_2, which is denoted by Q.

The schematic form in (60) allows us to distinguish between two kinds of quantifier-related binding, that of the variables $x_1,...,x_n$ and that of the variables $y_1,...,y_m$. We refer to the former kind as *primary quantificational binding* and to the latter as *secondary quantificational binding.* (We recall that the natural language examples of secondary quantificational binding we have seen so far all involve indefinite NPs in the restrictor of the quantifier represented by the central part of the duplex condition. The cases of secondary quantificational binding known to us are all of this kind.

The introduction of duplex conditions into DRT seems to bring its representation of quantification much more in line with that practiced in traditional logic than was the case for the original formulation of DRT. It should be emphasised however that the alternative possibilities of capturing quantificational effects which make DRT in its original formulation look so very different from standard formulations of predicate logic are still there. In fact, not only do we still have the quantificational interpretation of discourse referents in the universes of DRSs in the scope of $\Rightarrow$, $\neg$, and $\vee$, the duplex conditions themselves incorporate this alternative source of quantificational effects as well, viz. in the form of secondary quantificational binding. The point of including duplex conditions in the DRS formalism as an additional mode of representation is that the quantification expressed by those natural language constituents which duplex conditions are used to represent is fundamentally different from the quantificational effects that are produced by indefinites within the scope of operators like negation or conditionalisation. If these different forms of quantification seem to come to the same thing within the context of standard predicate logic, this should be seen as a symptom of the exclusively truth-conditional focus of predicate logic on the one hand and of its limited expressive resources on the other. Semanticists who are interested in truth conditions only will see this kind of simplification as harmless and maybe even as desirable. But it can be harmless only for so long as the quantifiers expressible within the formalism are those definable from the standard existential and universal quantifiers of the lower predicate calculus. As we have seen in this section, even truth conditions may be affected when non-standard quantifiers (such as *most*) are taken into consideration as well.

Duplex Conditions and Generalized Quantifier Theory

A large part of the more logically oriented literature on quantifiers is concerned with their formal properties (see [Westerstahl, 1989a; Keenan and Westerstahl, 1997; van der Does and van Eijck (eds.), 1991]). In particular, there is a long-standing concern to identify and study those properties which single out from the set of all logically possible generalised quantifiers those that are actually found in natural languages. Especially prominent among these properties is *conservativity*: A binary relation D between sets is *conservative* iff for any sets A and B, D(A,B) iff D(A,A $\cap$ B).

It is easily verified that all cases of quantification we have discussed so far (and which can be analysed as relations between sets) are conservative. In fact, conservativity is a consequence of the very way in which quantificational constructions are conceived in DRT. As first argued in Section 2.2 in connection with $\Rightarrow$, the antecedent of a conditional serves as context of interpretation for the consequent; and so the consequent of the conditional is to be seen as an addition of the information it explicitly provides to the information provided by the antecedent. Thus, if K_1 is the representation of the antecedent and K_2 the representation of the consequent in the context of the antecedent, the conditional comes (roughly) to the implication $\mathrm{K}_1 \Rightarrow \mathrm{K}_1 \uplus \mathrm{K}_2$. Much the same applies to the DRT analysis of quantificational NPs: the material that goes into the nuclear scope of the representing duplex condition is to be interpreted in the context of the restrictor, and the nuclear scope is to be understood as addition to the restrictor. The statements of the truth conditions for *most*-quantifications which was given in (51)–(57) are a direct continuation of this insofar as they take the form of relations between the sets S_1 and S_2, corresponding to the restrictor DRS K_1 and the extension K_2 of K_1 with the material from the scope. In other words, the truth conditions associated with a duplex condition of the form $\mathrm{K}_1 \underset{\mathrm{x}}{\overset{\mathrm{Q}}{\Diamond}} \mathrm{K}_2$ are of the form $\mathrm{F}_{\mathrm{Q}}(\mathrm{S}_{\mathrm{K}_1}, \mathrm{S}_{\mathrm{K}_1 \uplus \mathrm{K}_2})$ to begin with. Since evidently $\mathrm{S}_{\mathrm{K}_1 \uplus \mathrm{K}_2}$ is a subset of $\mathrm{S}_{\mathrm{K}_1}$, $\mathrm{F}_{\mathrm{Q}}(\mathrm{S}_{\mathrm{K}_1}, \mathrm{S}_{\mathrm{K}_1 \uplus \mathrm{K}_2}) = \mathrm{F}_{\mathrm{Q}}(\mathrm{S}_{\mathrm{K}_1}, (\mathrm{S}_{\mathrm{K}_1 \uplus \mathrm{K}_2} \cap \mathrm{S}_{\mathrm{K}_1}))$, the quantification represented by such a duplex condition is conservative by fiat.

Essentially the same is true for duplex conditions in which more than one discourse referent is subject to primary quantificational binding. "Conservativity" is now to be understood as a property of binary relations between n-place relations rather than sets, but the generalisation is obvious: let U be a given non-empty set, Q a 2-place relation between n-place relations over U — that is, $Q \subseteq \mathcal{P}(\mathrm{U}^n) \times \mathcal{P}(\mathrm{U}^n)$. Then Q is *conservative* if for all $A, B \in \mathcal{P}(\mathrm{U}^n)$, $\langle A, B\rangle \in Q$ iff $\langle A, A \cap B\rangle \in Q$.

In the literature on generalised quantifiers conservativity is only one of many quantifier properties of which the question has been raised whether all natural language quantifiers have them. But it is the only one about which the DRT analysis of natural language quantification carries immediate implications. As the discussion of other properties of quantifiers is not directly relevant from a DRT-centered perspective, this is not the place to pursue them further. We refer the reader to [Westerstahl, 1989a], and to the other publications cited there.

Before we conclude this discussion of quantifiers representable by means of duplex conditions, we must add an observation on what has come to be recognised as a general feature of quantification in natural language. The interpretation of natural language quantifiers often involves implicit restrictions in addition to the restrictions that are explicitly expressed in the sentence itself. (And with adverbial quantifiers, where sometimes no material within the sentence itself makes a contribution to the restrictor, the implicit restrictions will make up the restrictor on their own.) Following [von Fintel, 1994] and many others we represent the implicit restrictions on a given quantifier by means of an additional predicate C on the quantificationally bound discourse referents. (Thus C will in general be an n-place predicate, where n is the number of discourse referents $x_1, \ldots, x_n$ involved in primary quantificational binding.) Formally the implicit restriction takes the form of a supplementary condition "$C(x_1,\ldots,x_n)$" in the restrictor of the duplex condition, as represented in (60). Morevover, since C has to be resolved within the context in which the quantification restricted by it occurs, we take it to give raise to a presupposition, which is included in the initial (or "preliminary") representation as left-adjoined to the duplex condition which represents the quantification. (For the details of the DRT-based treatment that is assumed here see Section 4.) In general, this presupposition will also contain information that is relevant for the resolution of the predicate discourse referent C (to some particular predicate). Representing this higher order constraint as $P(\mathrm{C})$ we get (61) as representation for the contribution made by the quantifier of a sentence, instead of the slightly simpler form exemplified in (60).

(61)
$$\left\langle \left\{ \boxed{\begin{array}{c} \mathrm{C} \\ P(\mathrm{C}) \end{array}} \right\}, \boxed{ \boxed{\begin{array}{c} x_1 \ldots x_n \; y_1 \ldots y_m \\ \mathrm{C}_1(x_1,\ldots,x_n,y_1,\ldots,y_m) \\ \vdots \\ \mathrm{C}_r(x_1,\ldots,x_n,y_1,\ldots,y_m) \\ \mathrm{C}(x_1,\ldots,x_n) \end{array}} \; \underset{x_1 \ldots x_n}{\langle \mathrm{MOST} \rangle} \; \boxed{\begin{array}{c} z_1 \ldots z_k \\ \mathrm{D}_1(x_1,\ldots,y_m,z_1,\ldots,z_k) \\ \vdots \\ \mathrm{D}_s(x_1,\ldots,y_m,z_1,\ldots,z_k) \end{array}} } \right\rangle$$

As with many other cases of presuppostion the most difficult part of the theory of contextual restriction concerns the principles which govern the resolution of C. This is a problem about which we will say nothing here. We will turn to a certain aspect of this question in Section 3.5, where, as we announced already, we will consider frequency adverbs in their capacity of quantifiers over times.

Beyond Duplex Conditions

Many discussions of quantification in natural language suggest the implicit assumption that all quantifiers found in natural language are semantically like generalised quantifiers (viz. binary relations between sets, or, more generally, between n-place relations) — in our terminology, that natural quantifiers are generally to be represented in the form of duplex conditions, as schematically represented in (60) and (61). In the course of the past two decades, however, it has become increasingly apparent that this is not so: natural languages also have quantificational devices, many of them perfectly natural and even colloquial ones, which do not fit the generalised quantifier pattern [Keenan, 1992].

In the remainder of this section on quantification we discuss two examples of English quantifiers for which this is true. This is meant as a hint of how the representational approach of DRT may be extended to provide adequate representations for such forms of quantification, and also as a remainder of how much work still needs to be done in this area of natural language semantics, whether within a DRT-based framework or any other.

Our first pair of examples seems to resist representation by means of a duplex condition because it expresses a relation not between two sets, but between three.

(62) a. Not as many women as men were drinking orange juice.

b. More boys than girls got a present that made them happy.

(62.a) says that the set of men who drank orange juice is larger than the set of women who drank orange juice. One might want to argue that the actual quantifier involved in this example is a binary set relation which holds between two sets A and B iff A has larger cardinality than B, that this relation can be represented in the duplex format we already adopted, and that (62.a) differs from other sentences expressing this same relation only in terms of the mapping rules which lead from syntactic structure to this representation. But the difficulties which sentences like those in (62) present are in fact more serious. Consider (62.b), in which the pronoun *them* must be construed as referring to boys on the one hand and to girls on the other. The syntactic form of (62.b) — like that of (62.a) and other sentences in which a comparative construction occurs as part of a quantifying NP — suggests that at some level of semantic representation we must have a duplex-like structure with the content of the NP occurring to the left of the quantifier part and the sentence material that expresses the predicate of which this NP is an argument occurring to the right of it:

(63)
$$\boxed{\begin{array}{|c|}\hline x^{pl} \\ \text{“boys than girls”}(x) \\ \hline \end{array} \;\; \begin{array}{c} Q \\ x \end{array} \;\; \begin{array}{|c|}\hline y \\ \text{present}(y) \\ x \text{ got } y \\ y \text{ “made them happy”} \\ \hline \end{array}}$$

N.B. The superscript *pl* of the discourse referent x indicates that x originates from a morphologically plural NP and therefore can serve as antecedent for a plural pronoun like *they* or *them*; for discussion see Section 3.4. Furthermore we have adopted the practice of abbreviating parts of DRSs in the form of quasi-atomic DRS-conditions in which the predicate is given by an expression in scare quotes. To replace such a DRS by one that is fully worked out, these abridged conditions must be further expanded. Since the principles involved in those expansions do not matter to the point at issue, and paying attention to them would be likely to obscure it, we have decided that it is better to leave the conditions in question in the unfinished state in which they are presented. Henceforth we will proceed in this way whenever this seems expedient.

In (63) the pronoun *them* can be resolved to x if we assume that, as has been assumed for duplex conditions, the part to the left of the quantifier serves as context for the material to the right. At the same time, however, the restrictor part of (63) has to be processed in such a way that the two predicates λx.boy(x) and λx.girl(x) remain separable, so that each of them can be separately combined with the predicate in the nuclear scope. In fact, the quantifier and the comparative construction involved in the subject NP — severally represented, one might say, by the words *more* and *than* — arguably form a single semantic unit, and a single construction step should result in the four-component representation shown in (64.a). After further processing of the material in the nuclear scope (in the present example no further processing happens to be required for the material in the components left of the quantifier) this leads to (64.b). At this point two strategies seem possible. According to the first a further processing rule turns (64.b) into (64.c). The truth conditions for (64.c) are those of the set comparison quantifier MORE: the cardinality of its first agument exceeds that of its second argument. It is important to note that while this quantifier is of the binary set relation type distinctive of standard generalised quantifiers, it is not conservative. (For instance, if A has smaller cardinality than B but larger cardinality than A $\cap$ B, then MORE(A,B) is false, while MORE(A,A$\cap$B) is true.)

The second strategy is to analyse the *more* of (62) as denoting a 3-place operator, and accordingly to take (64.b) as the final semantic representation of (62.b). This 3-place operator would denote a 3-place relation MORE^3 between sets, such that MORE^3(A,B,C) is true iff —A $\cap$ C— ¿ —B $\cap$ C—. Analysed this way sentences like those in (62) can be seen to validate a certain form of conservativity: We have MORE^3(A,B,C) iff MORE^3(A,B, C $\cap$ (A $\cup$ B))). This observation is in keeping with the intuition that in a sentence like (62.b) the discourse referent introduced by the subject NP can serve as antecedent for the interpretation of the material which makes up the VP, just as we have found this to be the case for those quantifying subject NPs which give rise to duplex conditions of the form given in (61).

(64) a.
$$\begin{array}{|c|}\hline \left\{\begin{array}{c}\begin{array}{|c|}\hline x^{pl} \\ \text{boy}(x) \\ \hline\end{array} \\ \begin{array}{|c|}\hline x^{pl} \\ \text{girl}(x) \\ \hline\end{array}\end{array}\right\} \;\left\langle\begin{array}{c}\text{MORE} \\ x\end{array}\right\rangle\; \begin{array}{|c|}\hline y \\ \text{present}(y) \\ x \text{ got } y \\ y \text{ “made them hapy”} \\ \hline\end{array} \\ \hline\end{array}$$

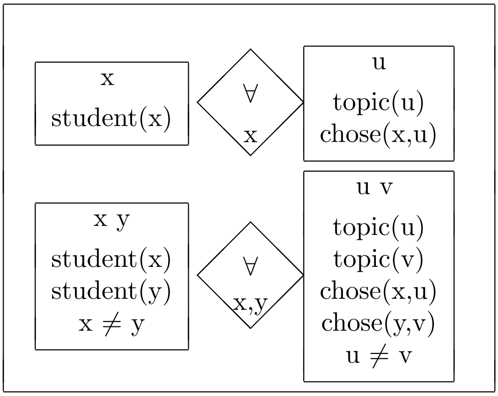
b.
$$\begin{array}{|c|}\hline \left\{\begin{array}{c}\begin{array}{|c|}\hline x^{pl} \\ \text{boy}(x) \\ \hline\end{array} \\ \begin{array}{|c|}\hline x^{pl} \\ \text{girl}(x) \\ \hline\end{array}\end{array}\right\} \;\left\langle\begin{array}{c}\text{MORE} \\ x\end{array}\right\rangle\; \begin{array}{|c|}\hline y\ u \\ \text{present}(y) \\ x \text{ got } y \\ u = x \\ \text{“make-happy”}(y,u) \\ \hline\end{array} \\ \hline\end{array}$$

c.
$$\left\langle\begin{array}{c}\text{MORE} \\ x\end{array}\right\rangle \left(\begin{array}{|c|}\hline x^{pl}\ y\ u \\ \text{boy}(x) \\ \text{present}(y) \\ x \text{ got } y \\ u = x \\ \text{“make-happy”}(y,u) \\ \hline\end{array}, \begin{array}{|c|}\hline x^{pl}\ y\ u \\ \text{girl}(x) \\ \text{present}(y) \\ x \text{ got } y \\ u = x \\ \text{“make-happy”}(y,u) \\ \hline\end{array}\right)$$

A quite different type of quantification which also does not fit the standard generalised quantifier pattern is that exemplified by (65)

(65) Every student chose a different topic.

(65) has a reading according to which it asserts (i) that every student chose a topic and (ii) that for every two different students x and y the topics chosen by x and y were distinct. It is obvious how the truth conditions associated with this reading should be written down in first order logic and, by the same token, how they can be represented in a DRS of the DRT formalism thus far developed. Such a DRS is shown in (66).

(66) 

(If it is assumed that the set of students has cardinality greater than 1, then the first duplex condition of (66) becomes redundant.)

As we found in connection with the sentences in (62), the challenge presented by sentence (65) is to explain how it is possible for the syntactic form of the sentence to give rise to truth conditions like those in (66). It is intuitively clear that the element of (65) which is responsible for the complexity of these truth conditions is the adjective *different*. But this observation doesn't help us much to account for how the subject NP of (65) and its object NP in which *different* occurs can "connive" to produce such truth conditions.

The intuition that these truth conditions are due to a coordinated contribution by the two NPs, and thus that these NPs are jointly responsible for a single, undecomposable quantificational complex at the level of logical form, is especially prominent in the work of Keenan on what he calls "non-Fregean quantifiers" (viz. [Keenan, 1992]). Keenan analyses sentences like (65) as involving a single quantificational operation on a 2-place relation.[18] The quantification is *polyadic* insofar as it binds two variables, corresponding to the two arguments of its relational operands. (In the case of (65) one of the variables corresponds to the subject NP and the other one to the object NP.) Schematically, the resulting logical form can be represented either as in (67.a), where O represents a function from 2-place relations to truth values and is applied to the 2-place relation $\lambda x.\lambda y.R(x,y)$, or as in (67.b), where O is a variable-binding operator on formulas which is applied to the formula R(x,y), while binding the two variables x and y. As vehicle for the truth conditions of sentences like (65) there is little to choose between those two forms and the little we have to say applies (mutatis mutandis) to either.

(67) a. $O(\lambda x.\lambda y.R(x,y))$

b. $(Ox,y).R(x,y)$

[18] A first treatment within DRT (and UDRT) can be found in [Seizmair, 1996].

In the case of (65) the formula in (67.b) is "student(x) $\wedge$ topic(y) $\wedge$ chose(x,y)" and the relation in (67.a) is the corresponding λ-abstract. Keenan shows that the operator involved in (65), schematically represented in (67.b) and yielding the truth conditions in (66), cannot be replaced by a pair of 1-place quantifiers which are applied to this relation one after the other. This entails in particular that we cannot associate an "ordinary quantifier" — which binds one variable only — with the subject NP of (65) and another such quantifier with the object NP in such a way that the successive application of these quantifiers to the mentioned formula (or alternatively to the formula chose(x,y) contributed by the verb on its own) lead to the truth conditions of (66).

While this is an important and interesting result, it doesn't solve the syntax-semantics interface problem to which we have drawn attention. It only underscores the urgency of that problem. As it stands we do not know how this problem should be solved, and we can only venture a speculation about the direction in which a solution might be found. Consider the sentence we get by eliminating *different* from (65):

(68) Every student chose a topic.

This sentence asserts the existence of a functional dependence of chosen topics on the students who chose them. There is substantial independent evidence (going back at last to the work of Engdahl [Engdahl, 1980] on functional wh-questions) that such functions can play a role in the interpretation of the sentences which either presuppose or entail their existence: Many sentences can be construed as making some claim about functions whose existence they entail or presuppose; and more often than not it doesn't seem possible to account for their meaning in another way.

Our tentative proposal for the analysis of (65) now comes to this. (65) entails the existence of the function we just described in connection with (68). (For it entails (68), which entails the existence of the function in its turn.) *different* (on the interpretation intended here) is to be construed as a predication of this function, to the effect that it returns different values for different arguments. However, in order that *different* can be applied to this function, the function has to be made available first. Thus, according to the present proposal the interpretation of (65) involves three distinct steps: (i) a "first run" interpretation in which *different* is ignored; (ii) the extraction of the function from this "first run" interpretation; and (iii) the application of *different* to this function. (It should be stressed that on this analysis relational adjectives like *different* act as predicates of functions *qua function*, i.e. as entities which embody information about what values the function returns for each of the arguments in its domain).

Even if this proposal should prove to be on the right track, it is evident that from the description we have given of it here many of the details are missing. The most serious shortcoming is that nothing has been said about the role that is played by the syntactic position which *different* occupies in (65), in virtue of which it can be interpreted as a predicate of the function which assigns topics to students. And how precisely does the relational character of the meaning of *different* contribute to the resulting reading?

The discussion of these last two examples has been speculative and the analyses we have suggested have many loose ends. We have included it nonetheless in order to make plain that the classical notion of generalised quantifiers as relations between sets is not the last word about quantification in natural language. The insight that natural language quantifiers very often work this way has been extremely important and fruitful, but it must not blind us to the fact that there is more. The same applies to duplex conditions. Duplex conditions constitute a non-trivial generalisation of the standard notion of generalised quantifiers as two-place relations between sets, but we have seen that they too cannot be applied in a straightforward manner to the analysis of the sentences in (62) and (65). These are but two examples of a range of cases of which we do not claim to fathom the diversity and complexity, but in all of which duplex conditions either are the wrong representational form or are related to syntactic structure via interpretation mechanisms that differ from those that are properly understood.

Metamathematical Properties of (Duplex Conditions for) Non-Standard Quantifiers

One often discussed type of question within the metamathematics of first order predicate logic and its various extensions is: what happens when we add one or more new quantifiers to a given formal language, and in particular what happens when we add these quantifiers to the classical first order predicate calculus itself? For instance: What can we say about the computational complexity of the notion of logical validity for the extensions of first order logic which result from such additions? Is validity still axiomatisable? If not, what is its complexity class (e.g. is it hyperarithmetic)? And what can be said about the extensions that are obtained in this way for certain natural subsystems of first order logic, such as monadic predicate logic?

Among the quantifiers with respect to which some of these questions have been explored we find in particular the quantifiers *most*, *there are infinitely many* and *there are non-denumerably many.* None of these are definable within standard first order logic. They form an interesting triad insofar as between them they exemplify the different answers that are possible for the question: what happens to axiomatisability when we add this undefinable quantifier to standard first order predicate logic? Briefly the answers are as follows. For the quantifier *infinitely many*: the extension is non-axiomatisable; for *non-denumerably many*: the extension is axiomatisable; for *most*: it depends. The first answer is an immediate consequence of Trakhtenbrot's Theorem (Non-axiomatisability of the Theory of Finite Models); the second is a famous early result in this area due to [Keisler, 1970]; the third needs explanation.

For *most* the situation is as follows. If we assume that the generalised quantifier MOST denoted by *most* satisfies the general condition: MOST(A,B) is true whenever $|A \cap B| > |A \setminus B|$ (i.e. irrespective of whether the sets A and B are finite or infinite), then the addition of MOST leads to non-axiomatisability. It is by no means evident, however, that this is the semantics for *most* that we should adopt. The condition $|A \cap B| > |A \setminus B|$ is plausible when A is finite, but far less so for cases where A is infinite. And alternative stipulations of the truth conditions of MOST, which arguably fit speakers' intuitions about what *most* means in the context of infinite sets better, can be shown to preserve axiomatisability.

We mention these few logical results about non-standard quantifiers because they illustrate what we consider an important point. Its importance will come more clearly into focus at the end of the next section. To prepare the ground for what we will say there we note the following. It appears that quantifying NPs, which have been the main topic of the present section always involve quantification over individuals, and not over sets. And the same appears to be true of adverbial quantifiers even if they sometimes involve quantification over several variables, rather than just one. This doesn't guarantee that adding such a quantifier to the first order predicate calculus will preserve its agreeable metamathematical properties, but it doesn't mean either that these properties will automatically be lost: for instance, Keisler's result [Keisler, 1970] shows that axiomatisability may be preserved even though the added quantifier is not definable, adding it therefore results in a genuine extension of standard first order logic. Whether a property such as axiomatizability will be preserved thus depends on individual features of the added quantifier.

Similar considerations apply to the addition of duplex conditions to the DRS-formalism of Section 3.1. From the addition as such nothing can be inferred about the metamathematical properties of the extension. Conclusions can be drawn only on the basis of the truth conditions associated with the particular quantifier symbols Q which occur in the central components of the added duplex conditions. We recall in this connection that all duplex conditions are "formally first order" in that the discourse referents involved (i.e. those subject to either primary or secondary binding) invariably stand for individuals, and not for sets. This restriction — that all discourse referents occurring in DRSs stand for individuals — will be abandoned in the next section.

3.4 Plural

Some of the quantifying NPs we discussed in the last section — such as *most tractors* or *all farmers* — were syntactically plural. But their semantic representation, it was stressed, always involved individual discourse referents — discourse referents whose values are individual farmers etc. Discourse referents standing for sets (of two or more elements) were not needed. When we consider definite and indefinite plural NPs this is no longer true.

The point is perhaps most easily made in connection with definite plurals such as *the farmers* or *the farmers of Weybridge*. The referents of such NPs must be represented as sets when predication — say by the verb of the sentence containing the NP — is collective. Thus

(69) The farmers of Weybridge voted against the by-pass.

has a prominent interpretation according to which the vote involved all the farmers of Weybridge and they voted against it as a body – some may have voted in favour but the majority was against and so the proposal didn't carry.

Such a predication can only be plausibly represented as a predication of the set consisting of the farmers of Weybridge. To this end we now introduce discourse referents representing sets (of cardinality ≥ 2) besides the ones we have been using so far, which always represent individuals. We use capitals for the new discourse referents, as opposed to the lower case letters which we continue to use for individual discourse referents. Thus the predication in (69) will take some such form as "X voted against", where X represents the set denoted by *the farmers of Weybridge*.

We do not want to pursue the analysis of definite NPs further at this point. We assume that all definites are presupposition triggers – they trigger presuppositions of proper reference. Accordingly their place is in Section 4, which is entirely devoted to presuppositional phenomena.

Indefinite NPs, however, are not presuppositional and plural indefinites resemble plural definites in that they can be the subjects of collective predication. Examples are the sentences in (70).

(70) a. Five/Some lawyers hired a new secretary.

b. Some graduates from Harvard Law School decided to set up a "legal clinic" for the poor of South Boston.

(70.a) can be understood as a joint hiring — the secretary will be working for the five lawyers — and (70.b) as saying that some group of Harvard graduates made a joint decision. Here too it is only by representing the indefinite NP via a "set" discourse referent that we can guarantee adequate representations of the collective predications. In particular, using the discourse referent X to represent the subject NP of (70.a) (and extending the construction algorithm in intuitively obvious ways on which we do not dwell here) we get for (70.a) the following DRS.

$$(71)\quad \boxed{\begin{array}{c} \text{X y} \\ \hline \text{lawyer*(X)} \\ |(\text{X})| = 5 \\ \text{secretary(y)} \\ \text{hired(X,y)} \end{array}}$$

N.B. The asterisk "*" turns the predicate of individuals that is expressed by a noun N into a predication N* of sets which is true of a set X if each member of X satisfies N. Thus lawyer*(X) is equivalent to the DRS condition

$$\boxed{\begin{array}{c} \text{x} \\ \hline \text{x} \in \text{X} \end{array}} \left\langle \begin{array}{c} \forall \\ \text{x} \end{array} \right\rangle \boxed{\begin{array}{c} \\ \hline \text{lawyer(x)} \end{array}}$$

and could be replaced by that condition if this was preferred.

It might be thought that when predications involving plural definites and indefinites are not collective, the contribution made by the NP to the semantic representation could in principle be accounted for through the exclusive use of individual discourse referents. However, this is often awkward — think of how to express the contribution of *five* to the sentence *Five lawyers voted for the proposal.* — and it lacks principled motivation. It seems clearly preferable to assume that plural definites and indefinites always introduce plural discourse referents, and to treat non-collective predications involving such NPs as the result of some operation of distribution over the represented set. (For details see [Kamp and Reyle, 1993].)

There exists a substantial literature on the semantics of plurals (for instance [Lasersohn, 1995] and [Winter, 2002]). Here we concentrate on the dynamic and trans-sentential aspects of the semantic contributions made by plural NPs, focussing in particular on plural indefinites and plural pronouns.

Plural anaphoric pronouns allow for interpretational strategies that are not found with singular pronouns. These strategies involve certain inferential principles that are needed to obtain the pronoun's antecedent. The initial goal of DRT's account of plurality was to identify these principles, and that will also be the main purpose of the present section.

Two of these inferential principles are illustrated in (72).

(72) a. Tom met Sue. They talked for quite a while.

b. Few boys of Lena's class showed up. They were smart.

Consider first (72.a). The construction of its DRS proceeds in the familiar left-to-right manner, with the representation of the first sentence providing the context for the second. If the construction is to run according to plan the

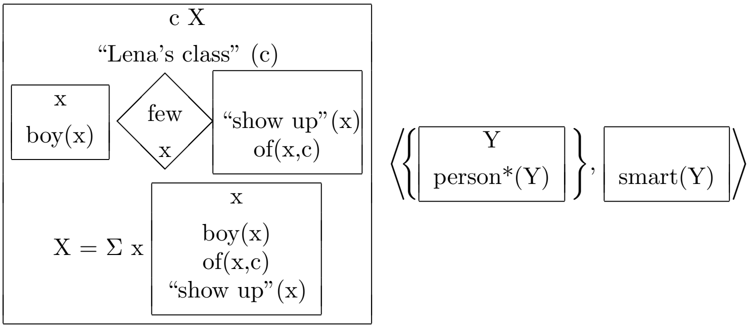
DRS for the first sentence should present a discourse referent that can serve as antecedent for the anaphoric pronoun *they*. But if the DRS construction for the first sentence follows the rules we have been assuming (for details see Section 2.3 or [Kamp and Reyle, 1993]), then the antecedent discourse referent which represents the set consisting of Tom and Sue is not available: no such discourse referent is a member of the DRS universe resulting from this construction; all that it contains are a discourse referent representing Tom and one representing Sue as shown by the DRS in the upper left corner of (73). In order to obtain the antecedent we want, we have to *synthesise* it out of what this DRS for the first sentence provides. The synthesisation operation which accomplishes this is called *Summation*. It takes a set Z of two or more discourse referents as input and returns a discourse referent representing the set consisting of all individuals represented by the different discourse referents belonging to Z. We represent applications of the Summation operation by adding the "output" discourse referent, say X, to the DRS universe where it is wanted while adding the condition "$X = \Sigma Z$" to the corresponding condition set. We use a capital letter for the new discourse referent since it invariably represents a set of two or more elements. In the case before us Z consists of the discourse referents t for Tom and s for Sue. In cases like this we write the condition "$X = t \oplus s$" instead of the official notation "$X = \Sigma\{t,s\}$". We assume in addition that Summation is applied as part of the effort to resolve the anaphoric presupposition that is triggered by the anaphoric pronoun *they*. Thus it is the combination of the completed DRS for the first sentence of (72.a) and the preliminary representation of the second sentence, in which the presupposition triggered by *they* is explicitly represented, that gives rise to the application in this instance. The result of applying the Summation operation is shown to the right of the first $\leadsto$; the DRS after the second $\leadsto$ results from the resolution of the anaphoric discourse referent Y to the Summation output X; this is the final DRS for (72.a).

(73)
$$\boxed{\begin{array}{c} \mathrm{t\ s} \\ \mathrm{Tom(t)} \\ \mathrm{Sue(s)} \\ \mathrm{meet(t,s)} \end{array}} \uplus \left\langle \left\{ \boxed{\begin{array}{c} \mathrm{Y} \\ \mathrm{person^*(Y)} \end{array}} \right\}, \boxed{\begin{array}{c} \\ \mathrm{talk(Y)} \end{array}} \right\rangle \leadsto$$

$$\boxed{\begin{array}{c} \mathrm{t\ s\ X} \\ \mathrm{Tom(t)} \\ \mathrm{Sue(s)} \\ \mathrm{X = t \oplus s} \\ \mathrm{meet(t,s)} \end{array}} \uplus \left\langle \left\{ \boxed{\begin{array}{c} \mathrm{Y} \\ \mathrm{person^*(Y)} \end{array}} \right\}, \boxed{\begin{array}{c} \\ \mathrm{talk(Y)} \end{array}} \right\rangle \leadsto \boxed{\begin{array}{c} \mathrm{t\ s\ X\ Y} \\ \mathrm{Tom(t)} \\ \mathrm{Sue(s)} \\ \mathrm{X = t \oplus s} \\ \mathrm{meet(t,s)} \\ \mathrm{Y = X} \\ \mathrm{talk(Y)} \end{array}}$$

The operation that is involved in providing the antecedent for *they* in (72.b) is called Abstraction. The Abstraction operation acts on duplex conditions and introduces a plural discourse referent X that stands for the set consisting of all individuals that satisfy the DRS K which results from merging the restrictor of the duplex condition with its nuclear scope. The DRS condition expressing this has the form X = Σx.K, where Σ is now to be understood as binding the discourse referent x. For the case at hand the condition is shown at the bottom of the DRS on the left in (74).

(74)



In addition to the interpretation represented in (74), (72.b) also has an interpretation in which *they* refers not just to the boys in Lena’s class who showed up, but to the set of all boys in Lena’s class. (In the present example this interpretation is awkward for rhetorical reasons, but it isn’t hard to come up with alternatives in which it is quite natural. For instance, we could replace the second sentence of (72.b) by *Nevertheless they had all received an invitation.*) A discourse referent representing this set can also be obtained through Abstraction. But in this case the operation has to be applied to the restrictor of the duplex condition on its own.

The examples in (72) have shown that certain inference-like operations — Summation and the two versions of Abstraction — may be used to synthesise the antecedents of plural pronouns from material present in the context DRS. Since antecedents may be derived from the context through the application of these operations, it might be thought that any logical derivation from the context DRS of the existence of a set may be used to interpret a plural pronoun. Since such inferences are generally not allowed for singular pronouns (see the discussion of (42)), the difference between plural and singular pronouns would thus simply come to this: the antecedent of a singular anaphoric pronoun must have been introduced explicitly by the DRS-construction algorithm (i.e. as the discourse referent representing some earlier NP); the antecedents of plural anaphoric pronouns may be logically inferred from the context in the form which the construction algorithm imposes on it.

This way of formulating the difference between singular and plural pronouns, however, is not only misleading, it is inadequate. To see this compare the following three sentence pairs:

(75) a. Two of the ten balls are not in the bag. They are under the sofa.
  b. Eight of the ten balls are in the bag. They are under the sofa.
  c. Few boys of Lena's class showed up. They were smart.

The *they* of (75.a) can be understood as referring to the two balls that are missing from the bag. In contrast, no such interpretation is possible for the *they* of (75.b). Nevertheless we can infer from the first sentence of (75.b) that there must be such a set — it is the difference between sets that are explicitly mentioned, viz. the set of eight balls that are in the bag and the larger set of ten balls of which this first set is said to be a subset. But, apparently, subtracting one set from another is not a permissible operation for the formation of pronominal antecedents. And so the inference to the existence of this set, while valid, is not sufficient to have it as antecedent for the pronoun. Similarly, a plural pronoun cannot pick up the complement of a group introduced by Abstraction. The *they* in (72.c) cannot refer to the group of boys that didn't show up.

Between them the sentence pairs in (72) and (75) show that the interpretation of anaphoric plural pronouns is supported by a restricted repertoire of logical operations which create pronoun antecedents from material in the context representation. These examples only give us a hint of what a precise characterisation of this repertoire could be like. We do not know that a formal definition of it has been attempted. But even without such a characterisation there are two conclusions that can be drawn, the first firm, the second more tentative.

The first is that what we are seeing here is a form of an aspect of linguistic knowledge: What operations may be used to construct antecedents for plural pronouns is an aspect of the interpretation of this particular type of expression. (Note in this connection that the restrictions we have observed in connection with (75.a,b) disappear when we replace the pronouns by definite descriptions such as *the missing/other balls* or *the boys who didn't come*.) Moreover, we are dealing with linguistic knowledge which pertains to the semantics of discourse, since it is often the discourse context, provided by the sentences which precede the one in which the given pronoun occurs, to which the antecedent-creating operations must be applied. (Especially for those who think of linguistic knowledge as confined to (syntactic) structure of the individual sentence, this is a phenomenon that ought to be food for thought.)

Secondly, what we have noted about the limited repertoire of logical operations available for the construction of pronoun antecedents suggests that the apparent difference between plural and singular pronouns to which we have drawn attention may well be reducible to the fact that singular pronouns stand for individuals while plural pronouns stand for sets of two or more individuals. We have noted that the operation of Summation always yields sets of cardinality ≥ 2. Moreover, this tends to be true of Abstraction as well, viz. in all those cases where the DRS K to which the operation is applied is satisfied by more than one value for the discourse referent bound by the abstraction operator. As a rule this last condition is fulfilled. (Often it is a presupposition associated with the linguistic construction which gave rise to the duplex condition from which K is obtained, e.g. nominal quantifiers such as *every boy*, *all/most boys*.) Thus application of either Summation or Abstraction will in the normal course of events produce discourse referents that are unsuitable antecedents for singular pronouns, even if we assume that nothing else speacks against their employment in the interpretation of such pronouns. Moreover, there are some cases where Abstraction does seem to be needed to interpret a singular pronoun, viz. where the sentence preceding the one containing the pronoun involves the quantifying phrase *there is exactly one ball* — as in (76).

(76) There is exactly one ball missing from the bag. It is under the sofa.

However, even if there is no difference here between the logical repertoire supporting the interpretation of plural pronouns from that of singular pronouns, it remains true that this repertoire is characteristic of the behaviour of pronouns, as one particular category of anaphoric expressions that we find in natural languages such as English.

Mereological vs. Set-Theoretical Ontology

In the DRSs displayed above graphically distinct discourse referents (lower case and upper case letters) have been used to represent single individuals and collections of two or more individuals. This could suggest that the graphically distinct discourse referents are meant to stand for entities of distinct ontological types, individuals and sets of individuals. However, in the model-theoretic semantics for the extended DRS formalism to which these DRSs belong no type distinction is made between the possible values of the two kinds of discourse referents. The ontology adopted in the models of this semantics is the mereological one first proposed for semantic purposes in [Link, 1983]. Link's proposal involves a single ontological category which provides for the denotations of mass terms (NPs whose nominal head is a mass noun) as well as singular and plural count terms (NPs with a count noun as head). In this survey we are concerned only with singular and plural count terms, so only that part of Link's ontology is relevant which concerns the denotations of those terms. This ontological category is structured by a part-whole relation $\leq$, and this part takes the form of an upper semi-lattice $\mathcal{A} = \langle \mathrm{A}, \leq \rangle$ which does not have a zero element (i.e. an element $0_{\mathcal{A}}$ such that for all $a \in \mathrm{A}$, $0_{\mathcal{A}} \leq a$) and which is complete, atomic and free.

DEFINITION 31.

(i) An *upper semilattice* $\langle \mathrm{A}, \leq \rangle$ is called *complete* if for all $\mathrm{X} \subseteq \mathrm{A}$ the supremum $\bigvee \mathrm{X}$ exists.

(ii) If a is the "largest" element of A — i.e. for all $x \in \mathrm{A}$, $x \leq a$ — then a is called the *one of* A and denoted as $1_{\mathcal{A}}$. Similarly, if a is the "smallest" element of A — i.e. for all $x \in \mathrm{A}$, $a \leq x$ – then a is called the *zero of* A and denoted as $0_{\mathcal{A}}$.

(iii) By an *atom of* $\mathcal{A}$ we understand any element $a \neq 0_{\mathcal{A}}$, such that $\forall x(x \leq a \rightarrow (x = a \vee x = 0_{\mathcal{A}}))$.

(iv) $\mathcal{A}$ is said to be *atomic* if for every $a, b \in \mathrm{A}$ such that $a \not\leq b$ there is an atom c such that $c \leq a$ and $c \not\leq b$.

(v) $\mathcal{A}$ is *free* if for all $a \in \mathrm{A}$, $\mathrm{X} \subseteq \mathrm{A}$ if $\mathrm{At}(a)$ and $a \leq \bigwedge \mathrm{X}$ then $(\exists b \in \mathrm{X})(a \leq b)$.

With respect to a model whose universe is such a lattice $\langle \mathrm{A}, \leq \rangle$ lower case discourse referents represent atomic elements of A and upper case discourse referents non-atomic elements. (Thus an assignment will have to map the lower case discourse referents onto atomic elements and the upper case discourse referents to non-atomic ones.) Moreover, the sum operation Σ and Abstraction operation Σx occurring within the new DRS conditions are interpreted as the join operation $\vee$ of the semi-lattice, while the *-operator gives the closure of predicates under $\vee$. I.e. if P is any 1-place predicate of our representation language whose extension is a subset V of A, then the extension of P* is the set of all $a \in \mathrm{A}$ which are joins of subsets of V.

Upper semi-lattices which are complete, atomic and free have a remarkably simple structure. It is easy to show (see, e.g., [Kamp and Reyle, 1993]) that such a structure $\mathcal{A}$ is isomorphic to a structure $\langle \mathcal{P}(\mathrm{B}), \subseteq \rangle$ where $\mathcal{P}(\mathrm{B})$ is the set of all subsets of some given set B and '$\subseteq$' is the relation of set-theoretical inclusion. In particular, one can take B to be the set $\mathrm{At}(\mathcal{A})$ consisting of all atoms of $\mathcal{A}$.

THEOREM 32. *Let $\mathcal{A} = \langle\, A, \leq \,\rangle$ be a complete, atomic, free upper semi-lattice without zero, and let At($\mathcal{A}$) be the set of atoms of $\mathcal{A}$. Then $\mathcal{A}$ is isomorphic to the structure $\langle \mathcal{P}(At(\mathcal{A})) \setminus \emptyset, \subseteq \rangle$.*

Theorem (32) shows that the choice between a lattice-theoretic and a set-theoretic approach towards the model theory of singular and plural count nouns is not important from a strictly formal point of view: models based on the one approach can be readily converted into equivalent models based on the other. Even so, there are considerations of naturalness which clearly favour the lattice-theoretic approach. First, the behaviour of singular and that of plural NPs are quite similar, both from a syntactic and from a semantic perspective. In view of this, making singular and plural NPs denote entities of logical types as different as individuals (i.e. first order entities) and sets of individuals (second order entities) seems to lack motivation.

A second, and more decisive, argument in favour of having a single entity type that includes the possible denotations of both singular and plural NPs is the following. Sometimes discourse referents must be allowed to take both "individuals" and "sets of individuals" as values. This can happen when a discourse referent is quantificationally bound. An example is provided by the sentences in (77).

(77) a. All boys bought the books they wanted.

b. All boys bought books they wanted.

(77.a) can be used to describe a situation in which some of the boys wanted a single book, and bought it, while the others wanted several books, and bought them. This means that in a DRS for (77.a) the discourse referent ζ introduced by the NP *the books they wanted* must be allowed to get as value a single book when the discourse referent x introduced by *all boys* and bound by the quantification denoted by *all* takes the first kind of boy, and a set of individuals when x gets mapped to a boy of the second kind. If we assume that ζ has a single logical type (an assumption usually made for variables of typed calculi), then its possible values must all belong to a single ontological category. The same applies to (77.b) in which the definite *the books* has been replaced by the indefinite *books.* We will return to this latter sentence below in a slightly different context, and will then also consider its semantic representation.

We conclude the present section with a brief statement of how the changes introduced in this section affect the model theory of the new extended DRS formalism. First, the model with respect to which DRSs of the extended formalism are to be evaluated are like those introduced in Section 3.1 except that the universe of a given model $\mathcal{M}$ now has the structure $\langle A, \leq \rangle$ of a complete, atomic, free join-semi-lattice. This extra structure is directly relevant only for the evaluation of atomic DRS-conditions. The satisfaction clauses for these conditions are given in the next definition. It is to be noted in this connection that the distinction between lower case and upper case discourse referents is treated as "syntactic sugar". In the official notation for the present DRS formalism there is still only one type of discourse referent, and discourse referents of this one sort can stand for any entity of the new ontology — i.e. for groups as well as for (atomic) individuals. We continue to use the old discourse referent symbols (i.e. x, y, z, ... x_1, x_2, x_3, ...) and distinguish between discourse referents which stand only for individuals, those which stand only for groups and those which allow for values of either kind by means of the predicate "at": a discourse referent x standing only for individuals comes with the condition "at(x)", a discourse referent x standing only for groups with the complex condition "$\neg\boxed{\text{at(x)}}$", and when neither of these conditions is present this means that x can take values of either kind. (Distinguishing between lower case, upper case and lower case Greek letters is nevertheless a useful practice. We have also found it convenient to refer to the first as "singular discourse referents", to the second as "plural discourse referents" and to the third as "neutral discourse referents".) In addition, we assume that the new formalism has the cardinality functor $|.|$ ($|x|$ is the "cardinality" of x, i.e. the cardinality of the set of atoms contained in x), and moreover, a canonical name n for each natural number n (e.g., the symbol "5" might be the canonical name for the natural number five.). Finally we need to be more explicit than we have so far been about the exact form of the condition introduced by applications of Summation and Abstraction. Conditions introduced by Summation have the form $\mathrm{X} = \Sigma\ \{\mathrm{y}_1,...\mathrm{y}_n\}$. These introduced by Abstraction take the form $\mathrm{X} = \Sigma\mathrm{y.K}$, where, we assume, y belongs to U_K.

Definition 33 extends Definition 10 of Section 3.1 with the new clauses that now become relevant. We now assume that the universe U of the model $\mathcal{M}$ has the structure of an atomic, free and complete upper semi-lattice $\langle\ \mathrm{U}, \leq\ \rangle$. (They clauses are listed starting with (ix) to make clear that we are dealing with an extension of Definition 10.)

DEFINITION 33.

(ix) $g \models_{\mathcal{M}} \mathrm{At(x)}$ iff $g(\mathrm{x})$ is an atomic element of $\mathrm{U}_{\mathcal{M}}$

(x) $g \models_{\mathcal{M}} \mathrm{x} \in \mathrm{y}$ iff $g(\mathrm{x})$ is an atomic element of $\mathrm{U}_{\mathcal{M}}$, $g(\mathrm{y})$ is a non-atomic element of $\mathrm{U}_{\mathcal{M}}$ and $\mathrm{x} \leq \mathrm{y}$

(xi) $g \models_{\mathcal{M}} \mathrm{x} = \Sigma\{\mathrm{y}_1,...\mathrm{y}_n\}$ iff $g(\mathrm{x}) = \bigvee\{\mathrm{b} \in \mathrm{U}_{\mathcal{M}} | (\exists \mathrm{y}_i)(\mathrm{b} \leq g(\mathrm{y}_i))\}$

(xii) $g \models_{\mathcal{M}} \mathrm{x} = \Sigma\ \mathrm{y.K}$ iff $g(\mathrm{x}) = \bigvee\{\mathrm{b} \in \mathrm{U}_{\mathcal{M}} | (\exists h)(g \subseteq_{\mathrm{U}_{\mathrm{K}}} h\ \wedge\ h(\mathrm{y}) = \mathrm{b}\ \wedge\ h \models_{\mathcal{M}} \gamma$ for all conditions $\gamma \in \mathrm{Con}_{\mathrm{K}}\}$.

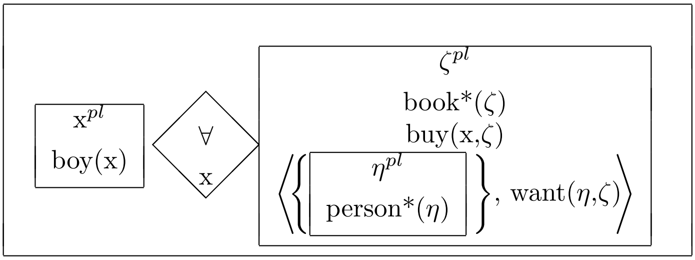
N.B. Both the model-theoretic extensions of the last subsection and the one stated in Definition 33 are easily adapted to the intensional set-up discussed in Section 3.2. We will return to the intensional perspective in the next Section, 3.5.

The Semantic Import of Plural Morphology

Most plural NPs we have considered in this section denote groups (of ≥ 2 members); they can be, and often must be, represented by plural discourse referents. In this respect they differ from quantificational NPs which, we argued in the last section, must be represented by singular discourse referents irrespective of whether they have singular or plural morphology. However, it is not only quantificational plural NPs where representation by means of a plural discourse referent is inappropriate. We saw that the direct object NPs *the books* and *books* of (77.a) and (77.b) require discourse referents that are neutral between individuals and groups. And in fact, (77) also gives us a further example of the need of singular discourse referents to represent plural NPs, viz. the plural pronoun *they*. The two occurrences of *they* in (77.a,b) can (and naturally would) be interpreted as anaphoric to the quantificationally bound singular discourse referent x introduced by the subject NP *all boys*. If the pronoun is interpreted in this way then the discourse referent that represents it must, through its identification with the singular discourse referent x, become a singular discourse referent – i.e. one whose values are restricted to (atomic) individuals — also; the discourse referent for the anaphoric expression inherits, so to speak, the features of its antecedent.

Note well that this anaphoric option for the plural pronoun *they* exists in (77.a,b) only because its intended grammatical antecedent, the subject NP *all boys*, is morphologically plural. When we change *all boys* into *every boy*, then (at least according to most English speakers we have consulted) *they* can no longer be interpreted as coreferential with the discourse referent bound by the quantifier (though it can be understood as referring to the set of boys as a whole; cf. the discussion of Abstraction earlier in this section). It thus appears that in each of the four instances of the phenomenon under discussion which we find in (77) — *the books*, *books* and the two ocurrences of *they* — there is some sort of dependency of the NP in question on another NP which is morphologically plural but whose semantic representation does not involve a plural discourse referent. This seems to be a general requirement. In the absence of such a dependency the discourse referent for a non-quantificational plural NP must stand for a group.

Let us assume that these dependency relations can be computed and then used to licence the interpretations that concern us here. This would make it possible to obtain for sentence (77.b) the representation given below in (79) via an initial representation like that in (78). In (78) the pronoun *they* is represented as part of an anaphoric presupposition; moreover, the discourse referent ζ for the NP *books* is given in a preliminary form, which creates the opportunity for its subject-dependent interpretation without yet establishing this interpretation.

(78)


New in (78) are the superscripts *pl* on the discourse referents x, ζ and η. A superscript *pl* indicates that the NP which is represented by a discourse referent bearing it is morphologically plural. Note further that in (78) the first of these superscripts applies to the singular discourse referent x, while in the other two cases it applies to the neutral discourse referents ζ and η. The explanation is this: the unequivocally quantificational character of the NP *all boys* determines that the discoure referent x which represents it must be a singular discourse referent, whereas the status of the discourse referents represented by the plural indefinite *books* and the plural pronoun *they* is initially undetermined.

Getting from (78) to (79) requires the application of two principles.

(i) The first principle concerns *pl*-marked discourse referents of anaphoric presuppositions triggered by pronouns. It says that an anaphoric *pl*-marked discourse referent may take a singular or neutral discourse referent as antecedent provided this antecedent discourse referent is also *pl*-marked.

(ii) The second principle concerns *pl*-marked discourse referents which are not to be anaphorically resolved. If such a discourse referent is neutral as it stands, and it stands in the right dependency relation to another discourse referent which is also *pl*-marked, then it may remain neutral in the final representation.

Applying the first principle in (78) to η^{pl} and the accessible discourse referent x^{pl} leads to identification of η with x. η is thereby coerced to act as singular discourse referent. There is no coercion of ζ so ζ retains its status of neutral discourse referent. Since the superscripts *pl* are no longer needed, they have been omitted in the final representation of (79). (In order to make the effects of the applications of the principles more clearly visible, we have used in the presentation of this representation the official notation introduced in the preceding subsection.)

(79)
$$\boxed{\begin{array}{c|c} \boxed{\begin{array}{c} x \\ \hline \text{boy}(x) \\ \text{at}(x) \end{array}} \;\Diamond_{x}^{\forall}\; & \boxed{\begin{array}{c} z\ y \\ \hline \text{book}^{*}(z) \\ \text{buy}(x,z) \\ y = x \\ \text{person}^{*}(y) \\ \text{want}(y,z) \end{array}} \end{array}}$$

What is missing from the way in which we have dealt with (77.b) is a proper account of the dependency relations that licence the two principles just mentioned. In fact, the constraints which govern dependent interpretations of plural definite and indefinite descriptions are not the same as those which govern dependent interpretations of pronouns. This is shown by the sentences in (80).

(80) a. Most boys have friends who have pets.

b. Most boys have a friend who has pets.

c. Most boys have friends they like.

d. Most boys have a friend they like.

In (80.a) we can understand *friends* as dependent on *most boys*, so that the sentence is neutral on the question whether the boys in question have one or more friends with pets, and *pets* can be interpreted as dependent on *friends*, so that each of the relevant friends could have had one pet or more than one. But for (80.b), where the "intermediate" NP *a friend* is in the singular, the only possible interpretation is that for each of a majority of boys there is a friend who has several pets. Here the plural morphology of *pets* and *most boys* does not licence a dependent interpretation of *pets*. Apparently, the relevant dependency relation which constrains dependent interpetations of indefinite descriptions as clause bound — the licencing plural NP must belong to the same clause as the licencee. For anaphoric pronouns the situation is different. In both (80.c) and (80.d) *they* can be interpreted as coreferring with the bound variable of the quantification. That in (80.d) there is a singular NP which "intervenes" does not seem to matter.

As regards plural pronouns, we conjecture that the constraint on applications of principle (ii) comes simply to this: in order that any discourse referent can serve as antecedent for any ananaphoric pronoun it must be accessible from the position of the pronoun (in the DRT-sense of accessibility). If moreover the discourse referent for the pronoun is *pl*-marked, then its antecedent must be either (a) a plural discourse referent, or (b) it must also be *pl*-marked. Thus the dependency constraint is in this case nothing more than the already familiar relation of accessibility. The dependency constraints involved in applications of principle (ii), on the other hand, do not seem to be reducible to notions we have already introduced. Apparently these constraints are at least in part syntactic. But in any case more needs to be said about them than we are able to do here and now.

The subject of this subsection has been one which belongs squarely within the realm of the syntax-semantic interface: how and why do certain sentence constituents, with their given morphological features and in their given syntactic positions, contribute to the semantic representation of the sentence in the ways they do? This might be thought a topic that is inappropriate in a survey article like the present one, where the focus ought to be on matters of logical and/or philosophical relevance. If we have decided to include the above discussion nevertheless, that has been for two reasons. First, syntax-semantics interface questions are of interest from a general philosophical perspective insofar as they reveal how complex the relationship between syntactic structure and logically transparent semantic representation can be: the principles according to which information is encoded in natural language differ significantly from those which determine the organisation of logical representation languages such as predicate logic or the DRS-languages of DRT. Exactly what these differences are and why they exist is surely of central interest for the philosophy of language, for the foundations of logic and for our general understanding of human information processing.

But there is also a second reason why we consider the problems that have been discussed in this subsection to be one of general interest. It is related to the first reason, but more specific. The general problem of the relation between meaning and linguistic form is often equated with the "syntax-semantics interface". Part of this equation is that syntax defines linguistic form and therewith both the input to and the constraints on the mapping which produces meanings as outputs. This is pretty much the standard view, and in first approximation it is surely correct. However, the problems we have discussed throw doubt on the simplicity of this view.

Apart from the dynamic aspects of this mapping, in view of which it is more appropriately seen and treated as a mapping from discourses and texts to meanings rather than from single sentences — this is the general message of dynamic semantics, and in no way specific to what has been said above — the discussions of this subsection have pointed towards the need for a "cascaded" mapping procedure, in which the syntactic structure of a sentence is first transformed into an "initial" semantic representation, and then from this initial representation into the representation which renders its semantic contribution in definitive, fully transparent form. However, it is clear neither of the initial representation itself nor of the operations that must be applied to it to turn it into the final representation whether they belong on the syntactic or the semantic side of the dichotomy that is implicit in a simple-minded conception of the relation between form and meaning. Particularly problematic in this context is the allocation of the operations

Summation and Abstraction. On the one hand these seem to belong on the side of meaning in that they operate on structures in which much of the information of the sentence has been made transparent already. On the other, they appear to be sensitive to aspects of linguistic form that are reflected in the "semantic" representations which serve as their application domains.

As we have seen intermittently in earlier sections, and will see in greater detail in Section 4below, the need for a cascaded architecture of DRS construction (and thus of the mapping from syntactic form to transparent representation of meaning) arises also for other reasons. The most important of these has to do with the treatment of presupposition which will be presented at length in Section 4. The basic assumption underlying this treatment is that presuppositions must first be given explicit representations which are then subjected to a (linguistically motiovated) process of presupposition justification, after which they either disappear from the semantic representation or else are integrated into its non-presuppositional core.

Metamathematical Properties

We conclude this subsection with an observation which links up with the concluding remarks of the last one. There we noted that the addition of non-standard quantifiers to the first-order DRS-formalism of Section 3.1 may but need not lead to the loss of axiomatizability. We also noted that the class of quantifying NPs include plural as well as singular forms. But all of these, we argued, introduce singular discourse referents. In contrast, indefinite and definite plural NPs, we have seen in the present section, are often (if not invariably) "semantically plural" in that they denote non-atomic entities (or, in more traditional terms, sets of cardinality ≥ 2). It is a consequence of this addition that the DRS formalism of this section necessarily transcends the boundaries of first order logic.

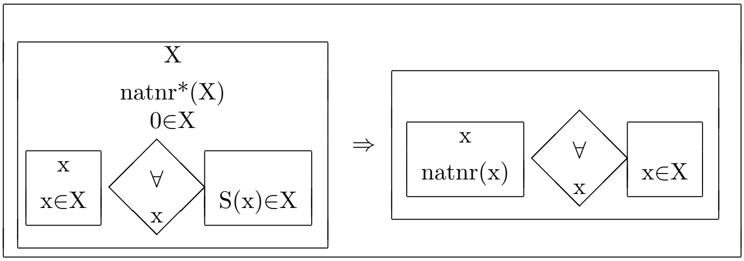
It should be stressed that it is the plural *indefinite* NPs that are the principal culprits here. To see this, observe that it is possible to state the induction axiom of second order Peano Arithmetic by means of the following English sentence:

(81) If 0 is among some numbers and the successor of a number is among them if that number itself is, then they include all the natural numbers.

(Note that this sentence does not make use of a noun such as "set". The crucial phrase, which is responsible for the irreducibly second order status of (81), is the NP *some numbers.*)

If we combine (81) with sentences which state the familiar axioms of Peano Arithmetic (those saying that the successor function is a bijection from the set of natural numbers without 0 to the set of all natural numbers, together with the recursive axioms for + and ×, then we get a version of Peano's second order axioms, an axiom system which has the standard model of arithmetic for its only model. This entails that the truths of arithmetic are an (easily recognisable) subset of the set of logical consequences of this axiom set. So, the set of these logical consequences has at least the complexity of that of the truths of arithmetic. Consequently it does not admit of a proof-theoretical characterisation. The same is true for a set of DRSs which give truth-conditionally correct representations of these axioms. In particular, if (82), which is a truth-conditionally correct representation of (81), is merged with DRSs for the other Peano axioms mentioned into a DRS K, the set of DRSs which are logically entailed by K is not amenable to proof-theoretic characterisation. Since all compounds of K other than (82) are first order (i.e. they can be given in a first order DRS language of the kind discussed in Section 3.1) non-axiomatisability must be due to (82), and thus to the presence in it of the plural discourse referent X, since that is the one feature of (82) which sets it apart from first order DRSs. And as far as the sentence in (81) is concerned, which (82) represents, the feature which makes it second order is the indefinite plural *some natural numbers* which is responsible for the presence of X in (82).

(82)



One noteworthy feature of this example is that it shows how little of the additional resources made available by the introduction of plural discourse referents is needed to move outside the realm of first order logic: (82) contains only one plural discourse referent, occuring within the scope of a single logical operator (viz ⇒). If the present formalism were restricted in such a way that an axiomatisable fragment results, hardly any of the additional expressive power that plural discourse referents introduce would be preserved.

It deserves to be stressed that it is an *indefinite plural* NP which is responsible for the second order status of (81). We saw at the end of the last section that non-axiomatisability may result from the incorporation of non-standard quantifiers. But whether this happens will depend on incidental and often subtle features of the particular quantifier in question. Moreover, since these quantifiers only bind variables ranging over (atomic) individuals, there is an important sense in which they do not transcend the bounds of first order logic. In this regard the extension of the present section is much more radical. It introduces a form of quantification over sets and this leads us directly into second order logic, with all the dire metamathematical concequences that entails.

3.5 Tense and Aspect

The starting point for DRT was an attempt in the late seventies to come to grips with certain problems in the theory of tense and aspect. In the sixties and early seventies formal research into the ways in which natural languages express temporal information had been dominated by temporal logic in the form in which it had been developed by Prior and others from the fifties onwards. By the middle of the seventies a large number of tense logics had been formulated, many of them for the very purpose of analysing temporal reference in natural language. It became increasingly clear, however, that there were aspects to the way in which natural languages handle temporal information which neither the original Priorean logics nor later modifications can handle. And some of these problems had to do directly with the behaviour of tense, i.e. with that feature of natural languages which had been a primary source of inspiration for the development of temporal logic in the first place. (In earlier days the term "tense logic" was the common way to refer to Prior-type temporal logics, and it is still used by many today.)

A particularly recalcitrant problem for the temporal logic approach are the differences between two past tenses of French, the imparfait and the passé simple, and the largely similar differences beween the past progressive and simple past in English. In many contexts it is possible to use either tense form to describe the same situation. An example is given by the pair of English sentences in (83).

(83) a. Hans was filling out his visa application form.

b. Hans filled out his visa application form.

On the face of it these sentences may seem to have the same truth conditions. But one feels that there is an important difference between them nevertheless. To explain what this difference consists in has been a problem of long standing. It was seen as a problem especially by those who taught French to non-native speakers and had to explain to their students when to use the passé simple and when the imparfait, and why.

Earlier attempts to account for the distinctions between these tenses (as well as between their English counterparts) were often couched in metaphorical or quasi-metaphorical terms. Thus, the differences between (83.a) and (83.b) have been variously described in terms like:

(84) a. (83.a) presents the event of which it speaks as "open", (83.b) presents it as "closed";

b. (83.a) presents its event "from the inside", (83.b) "from the outside";

c. (83.b) presents its event as "punctual", (83.a) as "temporally extended".

Moreover, it has been pointed out that when sentences in the passé simple or simple past occur in a narrative, they often "carry the story forwards", while sentences in the past progressive or imparfait hardly ever do this.

As they stand, these formulations are too informal and imprecise to be of much use in a systematic analysis. But they contain clear hints why it is that temporal logics aren't the right tools to explain what the difference between these tenses is. There are a number of reasons for this, three of which will be mentioned here. The first has to do with the temporal ontology on which an analysis of the tenses (and of temporal reference in natural language generally) should be based. Even a fairly cursory inspection of the way temporal reference works in natural language reveals that temporal intervals play as important a role as temporal points; in fact, from the perspective of natural language there does not seem to be a principled distinction between instants and intervals. This is at odds with Prior-type tense logics, with their commitment to the concept of "truth at an instant".

In the seventies alternative temporal logics — so-called interval logics — were developed to remedy this. But when we turn to the next two objections against temporal logics, interval logics are no real improvement on tense logics of the Priorean sort. The first of these has to do with the fact that in temporal logic — whether we are dealing with a point logic or an interval logic — time only enters at a meta-linguistic level. The formulas of temporal logics do not have any means to refer to times directly and explicitly — they have no terms whose values are points or intervals of time. This was originally seen as a virtue of tense logic. The tenses of the verb, it was thought, carry temporal information but they do this without making explicit reference to time. In this respect they are much like modal operators such as *necessarily* or *it is possible that*, which have to do with what might be or might have been the case as well as with what is, but do so without explicitly referring to possible worlds.

The principal reason why this is the wrong conception of the way in which temporal reference is handled in natural language is that tenses are not the only means that natural languages use for this purpose. As often as not the temporal information conveyed by a natural language sentence or discourse is the result of interaction beween several kinds of elements, of which the tenses constitute only one. Among the others we find in particular temporal adverbs and prepositional phrases such as *three hours later*, *on the first of February 2001*, etc. and these clearly do refer to times in a direct and explicit manner. (If these aren't explicitly referring expressions, what expression would count as referring?) Consequently, one would presume the logical forms, or semantic representations, for sentences in which such adverbs occur to contain devices for explicit reference to time as well. But the formulas of temporal logics do not. So they provide the wrong repertoire of logical forms.

A first hint of this is implicit in the informal observation concerning (83.a) and (83.b) that sentences in the simple past have the capacity to "carry the story forwards" while past progressive sentences do this hardly if ever. There are two sides to this difference — on the one hand simple past sentences differ from past progressives in the contexts which they contribute to the interpretation of the sentences which follow them; on the other there is a difference in the way in which simple past and past progressive sentences make use of the context which the sentences that precede them provide. Thus the dynamic dimension of what distinguishes these two tenses is both forward-directed and backward-directed.

The difference between the backward dynamics of simple past and past progressive comes roughly to this: When a past progressive sentence such as (83.a) occurs as part of a narrative passage or text, it is typically interpreted as describing a process that was going on at the time which the story had reached at that stage, i.e. as a process going on at the last time of the context established by the antecedent discourse. In other words, the temporal relation between the process described by the new sentence and the last time from the discourse context is that of temporal inclusion, with the process described by the new sentence including the time from the discourse context. In contrast, a simple past sentence like (83.b) is more naturally understood as presenting the event it describes as the next one in the sequence of narrated events, and thus as following the time reached thus far [Kamp, 1979; Kamp, 1981b; Kamp and Rohrer, 1983a; Partee, 1984].

We see this distinction between simple past and past progressive when we compare (85.a) and (85.b). Each of these is a "mini-text" consisting of two sentences. The only difference between them is that the second sentence of the first text has a past progressive whereas the second sentence of the second text is in the simple past:

(85) a. Josef turned around. The man was pulling his gun from its holster.

b. Josef turned around. The man pulled his gun from its holster.

The difference between (85.a) and (85.b) seems clear: in (85.a) the man is in the process of pulling his gun from its holster when Josef turns around and sees him. Here the second sentence is understood as describing a process that was going on at the point when Josef turned around. (85.b) is interpreted more naturally as saying that the event of the second sentence — i.e. that of the man pulling his gun from its holster — occurred after Josef turned around, presumably as a reaction to it.

An account of these differences in terms of the anaphoric properties of the tenses involved has to make use of some notion of context-supplied "reference point": the antecedent discourse provides a reference point (here as in many other cases: the time or event to which the story has so far advanced) with which the tense of the new sentence establishes a certain anaphoric relation. It seems natural therefore to build on what is undoubtedly the most famous early theory of tense in which the notion of reference point plays a prominent role, viz. that of [Reichenbach, 1947]. Unfortunately, however, Reichenbach's theory cannot be taken over as is. The difficulty has to do with what is arguably the most salient feature of this theory, its so-called "two-dimensionality". Reichenbach's theory is called a two-dimensonal theory of tense because it analyses the tenses in terms of pairs of relations, one between utterance time and what Reichenbach calls "reference time" and a second relation between reference time and the described "eventuality" (i.e. state or event).[19]

As we will argue below, Reichenbach's use of the notion of reference time suffers from the defect that it is "overloaded": in his theory reference times are made to do too many things at once. For this reason the DRT account of temporal reference has replaced Reichenbach's notion by a pair of notions which share the burdens of the original notion between them. They are: (i) the notion of "perspective time", which plays the role of Reichenbach's reference time in his analysis of the past perfect (about which more below) and which is responsible for the two-dimensionaltiy of the account presented here, and (ii) a second notion, for which the name "reference time" has been retained. It is the second notion which is used to account for the difference between (85.a) and (85.b).

To present this account in succinct terms is not all that easy. A number of preliminaries have to be dealt with before we can proceed towards the actual representations of (85.a) and (85.b), which are given in diagrams (87)–(89) below. Should the reader feel he is getting lost or bored, s/he might find it helpful to take a glance ahead at these diagrams.

For the time being we only consider those tenses for which temporal perspective time is not needed. (Simple past and past progressive, in the interpretation which matters in connection with (85.a), are among these.) The analysis of these tenses involves two relations, a relation between the utterance time and the described eventuality and a relation between this eventuality and the reference point. For both simple past and past progressive the first relation is that of temporal precedence — the eventuality precedes the utterance time. But with regard to the second relation the two tenses apparently differ. Simplifying somewhat: in the case of the past progressive the relation is temporal inclusion, in the case of the simple past

[19]We follow the widely adopted practice within the formal semantics of tense and aspect to use [Bach, 1981]'s term "eventuality" as a common term for the events, states, processes etc. which verbs can be used to describe.

it is not. (In the case of (85.b) the relation is temporal succession — the eventuality follows the reference time. In fact, the simple past often signifies temporal succession, but it doesn't always, a point to which we return below.)

The semantics of the tenses is complicated by a factor we have not yet mentioned. This is the role that is played by aspect. The term "aspect" covers a complex spectrum of interconnected phenomena. In the DRT-based theory of temporal reference we present here aspectual phenomena are considered only insofar as they have an effect on temporal reference. The theory assumes that it is possible to account for this influence by drawing a single binary distinction, that between events and states: Tensed clauses are assumed to have either stative or non-stative aspect; in the first case the eventuality described by the clause is a state, in the second it is an event. (So the totality of eventualities consists of two disjoint parts, the events and the states.)

In general the anaphoric properties of stative and non-stative clauses may differ even when they have identical tense morphology. This causes a problem for the line of analysis we have sketched, according to which it is the tense form alone which determines the temporal anaphoric properties of the clause. As it turns out, however, it is possible to deal with this problem without deviating too much from what we have outlined so far. All that is required is a slight complication of the analysis of the two above-mentioned relations – that between eventuality and utterance time and that between eventuality and reference time. The complication comes to this: we do not analyse the relations as directly involving the eventuality itself, but rather its "location time". Informally, the "location time" of an event is to be seen as the time when the event is said to occur and the location time of a state as the time at which the state is said to hold.

This leads to an analysis of tense in which the location time gets "interpolated", as it were, between eventuality on the one hand and utterance time and reference time on the other, and which involves three relations instead of two: (i) a relation between location time and utterance time; (ii) a relation between location time and reference time; and (iii) a relation between the eventuality and its location time. The difference between the simple past of the second sentence of (85.b) and the past progressive of the second sentence of (85.a) manifests itself through the third relation: In the case of the second sentence of (85.b) (or, for that matter the first sentence of (85.a) and (85.b)), the relation is inclusion of the described event in the location time. The past progressive determines the inverse inclusion relation: the location time is included in the described state.[20]

[20]This principle generalises to the progressive and non-progressive forms of other tenses, and beyond that to stative and non-stative clauses of any kind: when the clause is stative (and in particular when its verb is in the progressive), the location time is included in the described state; when the clause is non-stative, the described event is included in the

According to what we said so far, the semantic contribution of tense involves a combination of two things: on the one hand a relation between location time "t_{loc}" and utterance time "n" and on the other a relation between location time and reference time "r". The first contribution varies according to whether the tense is classified as past, present or future. The contributions are given in (86).

(86) Contribution of tense:

past $\boxed{\begin{array}{c} t_{loc} \\ t_{loc} \prec n \end{array}}$ *pres* $\boxed{\begin{array}{c} t_{loc} \\ t_{loc} = n \end{array}}$ *fut* $\boxed{\begin{array}{c} t_{loc} \\ n \prec t_{loc} \end{array}}$

The second contribution, we said, is of an anaphoric character. Adopting the position announced in Section 2, according to which anaphora is a presuppositional phenomenon, we analyse this contribution as taking the form of a presupposition: the tense of a clause triggers a presupposition to the effect that the location time of the eventuality which the clause describes stands in a temporal relation ρ to a reference time r; r has to be linked, via a process of anaphoric presuppposition resolution, to an element from the context. In the cases which concern us here this is the context established by the antecedent discourse.

Treating the second contribution of tense as a presupposition entails that the representation of a tensed clause involves two stages, one in which the presupposition is explicitly displayed and one in which it has been resolved. (Recall the sketch of this architecture in Section 2.3; for more details see Section 4 below.)

At last we come to the first of our representations. (87) gives a combination of the complete representation of the first sentence of (85.b) together with the preliminary representation of the second sentence. (For the first sentence of a discourse there is no discourse context. In such cases resolution of the tense-triggered presuppositions is governed by certain default rules; we ignore these here.) The first of the two relations contributed by the simple past of the second sentence — that between location time and utterance time — has been incorporated into the non-presuppositional part of the representation of this sentence (the structure on the right); the second relation is represented as presupposition prefixed to this non-presuppositional part (the structure in the middle). Note that the temporal relation ρ has not yet been identified with the one which we have argued gives the intuitively correct interpretation in this case, viz. that t_2 comes after r. The reason is that in general this relation depends on more factors than tense alone. More on this point below.

location time.

$$(87)\quad \begin{array}{|c|}\hline \mathrm{n\ j\ t_1\ e_1} \\ \text{Josef(j)} \\ \mathrm{t_1} \prec \mathrm{n} \\ \mathrm{e_1} \subseteq \mathrm{t_1} \\ \mathrm{e_1}\text{: “turn-around” (j)} \\ \hline \end{array} \quad \left\langle \left\{ \begin{array}{|c|}\hline \mathrm{r} \\ \rho(\mathrm{r},\mathrm{t_2}) \\ \hline \end{array} \right\}, \begin{array}{|c|}\hline \mathrm{n\ x\ t_2\ e_2} \\ \text{“the man” (x)} \\ \mathrm{t_2} \prec \mathrm{n} \\ \mathrm{e_2} \subseteq \mathrm{t_2} \\ \mathrm{e_2}\text{: “pull-gun” (h)} \\ \hline \end{array} \right\rangle$$

Before we say more about the presupposition of (87), a couple of remarks are in order about features of the representations shown in (87) which are fundamental to DRT's treatment of tense and aspect in general.

1. The first remark concerns the new types of discourse referents that are found in (87). Among these there are in the first place discourse referents standing for times (t_1, t_2, n). The presence of these discourse referents is enough to set the representation formalism which (87) exemplifies apart from the formalisms offered by the temporal logics mentioned at the start of this section, in which there is no explicit reference to time.

In addition we find in (87) also discourse referents for eventualities (e_1, e_2). After the informal discussion which led up to (87) this will hardly be surprising. Nevertheless it is a point which deserves special emphasis. The presence of these discourse referents is testimony to our conviction that (most) natural language sentences should be analysed as descriptions of eventualities. (More often than not the eventuality a sentence describes is introduced into the discourse by the sentence itself. But a sentence can also be used to provide an additional description of an eventuality that has been introduced previously.) Within semantics this view is at present hardly controversial and it is deeply embedded in the ways in which semanticists think about a wide variety of issues. From an ontological point of view however, events form a notoriously problematic category, for which identity conditions and structural properties and relations are hard to pin down and have proved a never ending topic for debate. For the philosopher and the philosophical logician it is therefore tempting to try and do without them.

There is a conflict here between philosophical conscientiousness and the needs of linguistic theory. (Cf. [Bach, 1981] on the distinction between "real metaphysics" and "natural language metaphysics".) We have argued elsewhere (see [Kamp and Reyle, 1993]) that this is a case where ontological theory just has to do the best it can, and that a form of "underspecified" model theory is the best option for realising this. A few words will be devoted to this in the last part of this section where we sketch the model theory for the DRS language to which (87) and the following DRSs of this section belong.[21]

[21] Of the arguments in favour of the view that verbs are to be treated as predicates of events, processes or states we mention here just two:

(i) This argument has to do with the relation between deverbal nouns and the verbs

2. The second remark concerns the occurrences of the discourse referent n. n is an *indexical discourse referent.* It represents the utterance time of the represented discourse or sentence. ("n" stands for "now".) Its presence introduces a pragmatic element into our representations which played no role up to this point. DRSs which display occurrences of n have to be understood not as representations of sentence or discourse *types*, whose identity is determined exclusively by linguistic form, but as representations of particular utterances (i.e. sentence or discourse *tokens*), which are made at some particular time.[22] The pragmatic element which n introduces into a DRS has implications for verification and truth. The verification of a DRS K which contains occurrencs of n like those we find in (87) must take account of the time t_u at which the represented sentence or discourse is uttered. This means that only assignments f are to be considered for which $f(n) = t_u$. A consequence of this is that verification of K is only possible in models (and, in case we are dealing with intensional models, at worlds of those models) in which the time of the utterance represented by K exists. (For more on this check the model theory for the DRS language which is presented in the last part of this section.)

With these two general observations behind us we return to the details of example (85.b). (87), we pointed out, contains a presuppositional component — viz. the DRS in curly brackets. Resolution of this presupposition requires finding specifications of two of its constituents, r and ρ. For sentences in the simple past the relation ρ is, as we noted in our informal gloss on the sentences in (85) above, often that of temporal succession: the new event e_2 occurred after the last event reached by the discourse so far, here e_1. To convert (87) into a representation which expresses this relation, the

from which they are derived. It is generally assumed that deverbal nouns (such as *walk*, *cleaning* or *action*, etc.) are treated as predicates, and thus, inevitably, as predicates of some such entities as events. (This, to our knowledge, is an assumption that has never been seriously challenged.) If that is right, however, then it is highly artificial not to treat the verbs from which these nouns are derived as event predicates also.

(ii) Treating verbs as predicates of eventualities restores an apparently universal feature of natural language predication which is lost when verbs are denied an eventuality argument. Note that all predicate-like word classes of natural languages other than verbs — in particular nouns, adjectives/adverbs and prepositions – are analysed as predicates one argument of which is syntactically implicit: This argument is not realised by a separate phrase, but carried by the lexical predicate itself. Thus a "relational" noun such as *friend* has one "internal" argument, which can be realised by an adjoined *of*-PP, such as *of Maria* in the NP *a friend of Maria.* But the other argument, the one for the person who is Maria's friend, cannot be realised by an explicit argument phrase. We find the same with non-relational nouns such as *girl*, *broom*, etc. Here there is only one argument and it is implicit in the same way as one of the two arguments of *friend.*

Note well that both these considerations apply as much to stative as to non-stative verbs.

[22] As well as by some particular speaker, and usually addressed to some particular person or audience, but these last two factors are of no importance in this section and will be left aside.

anaphoric discourse referent r must be resolved in the context that is provided by the DRS for the first sentence of (85.b). For all we have said so far, r could be resolved either to e_1 itself or to its location time t_1. We adopt the second option without argument.

We will return to the problems connected with the presupposition of (87) below, at the point when we will be in a position to compare the analysis of (85.b) with that of (85.a).

After resolving r to t_1, specifying ρ as $\prec$, incorporating the (now justified) presupposition into the representation of the second sentence and then merging the two DRSs into one, we get as final representation for (85.b) the structure given in (88).

(88)
$$\boxed{\begin{array}{c} n\ j\ t_1\ e_1\ x\ t_2\ e_2 \\ \text{Josef}(j) \\ t_1 \prec n \\ e_1 \subseteq t_1 \\ e_1\text{: “turn-around”}(j) \\ t_1 \prec t_2 \\ \text{“the man”}(x) \\ t_2 \prec n \\ e_2 \subseteq t_2 \\ e_2\text{: “pull-gun”}(x) \end{array}}$$

We have seen that (85.a) differs from (85.b) in that the eventuality of the second sentence is now understood as a process that is going on at the time of the event of the first sentence, and not as an event which follows it. Given the way in which we have analysed the second sentence of (85.b) in (87) and (88) one would expect this difference between (85.a) and (85.b) to manifest itself as a difference in the relation ρ: ρ should now be inclusion. More precisely, assuming that r is once more resolved to t_1, the relational condition of the presupposition should now be "$t_1 \subseteq t_2$".

This comes fairly close to what we want, but as it is, it isn't quite right. It isn't for two separate reasons. The first is that what we are really after is the conclusion that the eventuality e_1 described in the first sentence is included in the eventuality described in the second. Given the assumptions we have made, the condition "$t_1 \subseteq t_2$" does not guarantee this; it would if we could also rely on the condition "$t_2 \subseteq e_2$". But that is not what (87) tells us: It specifies that "$e_2 \subseteq t_2$", not that "$t_2 \subseteq e_2$".

The second reason is that even if the relation between e_2 and t_2 is reversed, there is still a problem with the characterisation of e_2 by means of the condition "e_2:"pull-gun"(x)". In the literature on tense and aspect this second problem is known as the *imperfective paradox* (see [Dowty, 1979]). It manifests itself in connection with a number of different linguistic constructions: with progressive forms of event verbs in English, with imparfait sentences in French, and also in a number of other situations which are of no direct concern here. When an event verb occurs in one of these constructions it is usually interpreted as describing a process that can be viewed as an initial segment of an event of the kind that the verb describes when it is used "non-progressively". However, there is no requirement that the described event segment actually evolves into a complete event of this kind.[23] The imperfective paradox is exemplified by (85.a) in that this sentence can be used to describe a scene in which the man never managed to get his gun out of its holster because Josef, with his widely known and much-feared reflexes, made sure he didn't.

The DRT account we present here deals with these two problems by not treating the past progressive as a complex tense (as was implied by the discussion up to this point), but by assuming, rather, that the past progressive factors into (i) the past tense; and (ii) an aspectual operator PROG. PROG transforms an event type E into a state type PROG(E). In particular, when the event predicate in the scope of the progressive is (as it is in our example) given by the event typing condition "e: "pull-gun" (x)" (see the right hand side representation of (87)), then the condition which characterises the state s as being of the corresponding PROG-type has the form "s: PROG($^\wedge$e.e: "pull-gun" (x))". (Here $^\wedge$ is the intensional abstraction operator of Intensional Logic, see Section 3.2, Definition 30.)

The past tense of a verb occurring with past progressive morphology is assumed to carry the same semantic import as the past tense of a sentence in the simple past (such as, e.g., the second sentence of (85.b)). Thus the contribution which the past tense makes to the semantics of the second sentence of (85.a) is the same as the one it makes to the three other sentences of (85.a,b). In all cases it is the contribution given by the past tense specification in (86). The difference beween the interpretation of the second sentence of (85.a) and that of (85.b) now results from an aspectual difference between events and states. As already noted an event is assumed to occur within its location time, whereas a state is assumed to be going on during its location time.

With these additional assumptions we get for (85.a) the representations (89.a) (corresponding to (87)) and (89.b) (corresponding to (88)).

$$(89)\quad \text{a.}\ \begin{array}{|c|}\hline \mathrm{n\ j\ t_1\ e_1} \\ \mathrm{Josef(j)} \\ \mathrm{t_1 \prec n} \\ \mathrm{e_1 \subseteq t_1} \\ \mathrm{e_1\text{: "turn-around" (j)}} \\ \hline \end{array} \left\langle \left\{ \begin{array}{|c|}\hline \mathrm{r} \\ \rho(\mathrm{r,t_2}) \\ \hline \end{array} \right\}, \right\rangle \begin{array}{|c|}\hline \mathrm{n\ x\ t_2\ s_2} \\ \text{"the man" (x)} \\ \mathrm{t_2 \prec n} \\ \mathrm{t_2 \subseteq s_2} \\ \mathrm{s_2\text{:PROG(}^\wedge e_2\text{: "pull-gun" (x))}} \\ \hline \end{array}$$

[23] Although there does appear to be a requirement that a completion into such an event be intended by the agent, or that the segment would have turned into one if it hadn't been for some external interference which prevented this.

b.

n j t_1 e_1 x t_2 s_2
Josef(j)
$t_1 \prec n$
$e_1 \subseteq t_1$
e_1: "turn-around" (j)
$t_1 \subseteq t_2$
"the man" (x)
$t_2 \prec n$
$t_2 \subseteq s_2$
s_2:PROG($^\wedge e_2$: "pull-gun" (x))

Discussion: How to choose the Reference Time, Temporal Relations and Discourse Relations?

The analysis of (85.a) and (85.b) we have shown here has been challenged by linguists who share many of our assumptions [Partee, 1984; Hinrichs, 1986; Roßdeutscher, 2000]. The controversy concerns the interpretation of the two elements of the tense-triggered presupposition — see (87) — which need resolution in context, viz (i) the reference time r and (ii) the relation ρ in which r stands to the new location time. The mentioned authors argue for a different conception of narrative progression, and with that for a different way of identifying reference times. On this alternative view an event sentence in a narrative introduces not only the event it describes into the discourse context but also a "reference point" which follows this event and acts as the (default) location time for the eventuality of the next sentence. Stative sentences do not introduce such a subsequent point. They inherit their "reference point" from the context in which they are interpreted and pass it on to the next sentence. This is one reason why, on the account now under discussion, event sentences propel the story forward but stative sentences do not. On this alternative account the determination of the relation ρ becomes simpler: ρ is always identity between the reference point and the new location time.[24]

The simple and uniform way in which the alternative account handles ρ seems to speak in its favour. But on the other hand it also encounters certain difficulties which do not arise for the approach we have presented. For instance, it cannot explain directly why (85.a) seems to imply that the pulling of the gun out of its holster was going on when the event e_1 of Josef turning around occurred. All it yields is that the process goes on at the reference point following e_1; that it was going on at the time of e_1 itself must be attributed to some further inference. It is clear that a motivated choice between the two accounts requires looking at many more examples

[24]Rossdeutscher's account differs from those of Hinrichs and Partee in ways that are difficult to explain at this point. We return to this later.

than the pair that has been considered here. However, when one looks more closely at examples that might help to decide between the two accounts, one finds that the crucial judgements not only tend to be delicate and unstable, but also that they are influenced by factors that neither account takes into consideration. What is really needed, is therefore not so much a choice between these two accounts, but a theory which is capable of dealing also with these additional factors.

Many of these factors have to do with rhetorical and other discourse relations.[25] (90) gives a few simple examples in which the effects of rhetorical relations on temporal relations are easy to perceive. In (90.a) the event reported in the second sentence may overlap the one reported in the first, $e_1 \bigcirc e_2$. This is typically the case if the second sentence is an *elaboration* of the first. A reversed temporal order, $e_2 \prec e_1$, is induced by the causal relation between the events reported in (90.b), where the second sentence is understood as giving a (causal) *explanation* of the first. And in (90.c) temporal progession, $e_1 \prec s_2$, (instead of overlap) is also induced by an assumption of causality. Here the second sentence issues to describe a *result* of the event reported in the first rather than a state that obtained while the event occurred.

(90) a. Chris had a fantastic meal. He ate salmon.

b. Max fell. John pushed him.

c. John turned off the light. The room was pitch dark.

Theories of discourse interpretation [Moens and Steedman, 1988; Hobbs, 1990; Caenepeel, 1989; Lascarides and Asher, 1993; Asher, 1993] use rhetorical and other discourse relations to represent the conceptual glue between the eventualities reported. The first theory to deal with discourse relations within a formal dynamic setting was Asher's S(egmented) D(iscourse) R(epresentation) T(heory). (See [Asher and Lascarides, 2003] as well as the SDRT publications cited there. Another important body of work in this area is that of Webber and others. See [Webber, 1988; Stone, 1998].) SDRT exploits non-monotonic logic to determine the possible interactions between discourse structure and temporal structure. Updates trigger usually defeasible inferences from an axiomatic system combining discourse

[25] One way to get a sense of the different factors that affect our judgments of temporal relations in discourse is to try to construct minimal pairs like that of (85), where the only difference is that a certain verb occurs in one member of the pair in the simple past and in the past progressive in the other. In order for it to be easy to get the contrast which (85) is meant to illustrate, the verb in question must allow on the one hand an interpretation according to which the action it describes is something which the agent could have been engaged in independently of (and thus antecedently to) the event described in the immediately preceding sentence, and on the other hand as an action which the agent could be seen as performing as a reaction to that event. When either of these requirements fails, one of the two texts becomes infelicitous or the intended contrast is no longer salient.

relations, temporal relations and world knowledge. Cases where different sources supply conflicting conclusions about interpretation are dealt with by the underlying non-monotonic logic.

We wish to stress, however, that it is nevertheless important for a theory of temporal interpretation which accounts for the correlations between temporal relations and discourse relations to also pay due attention to all constraints that are imposed on temporal relations by linguistic form. We refer to [Roßdeutscher and Reyle, 2000]. The strategy adopted there is in essence the same that is implicitly assumed in this entire chapter: First, an interpretation is constructed on the basis only of the linguistic information contained in the interpreted sentence or sentences. The temporal relations contained in the representation which results from this first, "purely linguistic" interpretation process will often be underspecified. However, further interpretational operations, which use the initial representation as input, may compute the discourse relations between the represented sentences. On the strength of these discourse relations the initially underspecified temporal relations may then be resolved or the underspecification reduced. Overall, this strategy does not differ essentially from SDRT in its current form [Asher and Lascarides, 2003].

Tenses and Temporal Adverbs

So far we have only considered sentences for which the interpretation of tense involved a link to a reference time supplied by the antecedent discourse. Through this link between location time and reference time the new eventuality is temporally located in relation to the context established by the preceding sentences of the discourse of which the new sentence is an integral part. It is just as common, however, for the eventuality to get temporally located sentence-internally, through the presence of a temporal adverb. A few examples of such cases of sentence-internal location are given in (91)

$$(91)\quad \left.\begin{array}{l}\text{On 03 03 03}\\ \text{Once upon a time}\\ \text{On the last day before his marriage}\\ \text{On a Sunday}\\ \text{On the preceding Sunday}\\ \text{The next day}\\ \text{Last Sunday}\\ \text{Yesterday}\\ \text{Yesterday, between 4.00 and 6.00}\\ \text{After the exam}\\ \text{During the summer holidays}\end{array}\right\}\ \text{Fred bought a lawn mower.}$$

Each of these different adverbs gives information about the time when Fred bought a lawn mower.

The first point to notice about the examples in (91) is the variety of referential mechanisms involved in determining what times the adverbs denote. Temporal adverbs display the full range of referential possibilities which we find with noun phrases in general – absolute, anaphoric, indexical, etc. This of course is no surprise, given that many temporal adverbs have the form of prepositional phrases. In addition, we find temporal adverbs subject to referential mechanisms which depend crucially for the way they work on the structure which we ascribe to time — the fact that time is a linearly ordered medium, with a metric grid imposed on it by the accepted calendar (manifest in our language through our ways of referring to particular times and dates, often with the help of calendar-related predicates like *day*, *week*, *month*, *year*, ...) An example of one such mode of specifically temporal reference is the possibility of using the phrase *on Sunday* to refer to the last Sunday before the utterance time, or alternatively to the first Sunday after it (with the tense of the sentence usually disambiguating between these alternatives). A systematic study of the range of referential possibilities for temporal adverbs is instructive (as well as indispensible for practical needs of computational linguistics), but it is not a matter we pursue here.

The second point is one we need to consider more closely. It concerns the way in which the referent of a temporal adverbial adjunct gets connected with the information provided by the rest of its clause. There are two aspects to this question. First, there is an issue of the syntax-semantics interface: how does the syntactic relation in which the temporal adverb stands to the remainder of the clause lead to its interpretation as temporal location predicate of the described eventuality? It is generally assumed that adverbial phrases are adjuncts, though there appears to be some degree of uncertainty about where such phrases are adjoined. But these details need not detain us, as long as we assume (i) that the constituent to which the adverbial is adjoined acts as a predicate of a certain argument and that the temporal locating adverb provides an additional predication of that argument, and (ii) that when the adjunction is to some syntactic projection of the verb, as it is in the sentences of (91), then this argument is the eventuality described by the verb. (Temporal adverbials aren't always adjoined to a projection of the verb. For instance, in a construction such as *the news at 12.00* the PP *at 12.00* is adjoined to the NP *the news*. Here, the argument of the locating predication is the referent of this NP. However, in the remainder of this discussion we limit attention to the cases where the adverb serves to locate the eventuality described by the verb.)

The other aspect of the contribution which temporal adverbs make to eventuality location has to do with content and form of the predications that temporal locating adverbials express. In DRT-terms: what are the discourse referents and conditions which the adverb contributes to the DRS of its clause? Since the sentence DRSs to which we have already committed ourselves involve not only eventualities but also their location times, the first question we need to answer here is: should the adverb be construed as locating the eventuality by entering in relation to it directly, or does it do this via a relation with its location time? It is not easy to motivate an answer to this question. We have adopted the second option. For the somewhat complicated and partly theory-internal reasons for this decision see [Reyle *et al.*, 2003]).

The contribution of a temporal locating adverb, then, takes the form of some relation between (i) its own referent and (ii) the location time of the described eventuality. To explore how this relation should be represented we need to look at some particular cases. It is advisable to begin with an adverb which does not have the form of a prepositional phrase; our choice is *yesterday*.

First some details concerning this particular adverb. *yesterday* is a deictic adverb. Normally it refers to the day preceding that on which the utterance containing it is made. (In special cases of indirect discourse — especially of so-called free indirect discourse — it may refer to the day preceding some past vantage point, but these we ignore; but compare the related remarks on shifted *now* in the next subsection.) We will represent this indexical information by introducing a discourse referent t_y to represent this day, together with certain conditions which determine how this day is determined in relation to n. To this end we make use of a partial functor DAY-OF which maps any time t that is included within some calendar day onto that calendar day, as well as of a predicate DAY which is true of those and only those periods of time which are calendar days. (Exactly what people understand by "day" may be open to some variation. For simplicity we assume that a calendar day runs from midnight to midnight.)

So much for the particularities of *yesterday* as distinct from other temporal adverbs. What is still missing is the relation between the discourse referent t_y which represents its referent and the discourse referent t_{loc} for the location time of the eventuality. Let us focus on the sentence in (91) which begins with *yesterday*. (That is, the sentence *Yesterday Fred bought a lawn mower.*) In this case we obtain an intuitively correct representation of the truth conditions of the sentence if we assume that the relation is that of inclusion of the location time within the adverb time: "$\mathrm{t}_{loc} \subseteq \mathrm{t}_y$". Together with what has already been assumed about the representation of tensed sentences and the special conditions connected with the reference of *yesterday* this condition leads to the representation in (92)

(92)
$$\begin{array}{|c|}\hline \mathrm{n\ f\ z\ e\ t}_{loc}\ \mathrm{t}_y \\ \mathrm{Fred(f)} \\ \mathrm{lawn\ mower(z)} \\ \mathrm{t}_{loc} \prec \mathrm{n} \\ \mathrm{DAY}(\mathrm{t}_y) \\ \mathrm{t}_y \supset\!\subset \mathrm{DAY\text{-}OF(n)} \\ \mathrm{t}_{loc} \subseteq \mathrm{t}_y \\ \mathrm{e} \subseteq \mathrm{t}_{loc} \\ \mathrm{e{:}buy(f,z)} \\ \hline \end{array}$$

(The symbol $\supset\!\subset$ denotes the relation of "abutment". An interval t_1 abuts an interval t_2 if (i) t_1 lies entirely before t_2, and (ii) there is no interval t_3 such that t_1 lies entirely before t_3 and t_3 lies entirely before t_2.)

Is the condition "$t_{loc} \subseteq t_y$" incidental to this particular sentence, in which a described event is located by the adverb *yesterday*? The answer is no. "$t_{loc} \subseteq t_y$" is the condition which links the referent of the temporal locating adverb to the location time in all cases where the adverbial is adjoined to a projection of the verb. This claim may well seem counterintuitive and we need to consider a couple of other cases to show why it can be upheld.

Adverbials which look at first blush like counterexamples are PPs begining with the prepositions *before* and *after*. Consider the sentence *After the exam Fred bought a lawn mower.* Isn't the relation between t_{loc} and t_{adv} temporal succession in this case, rather than inclusion? The reason why we maintain that this is not so rests on a view of the semantics of locating PPs which has been proposed in connection with adverbials of spatial location, but which, we believe, applies equally to those which locate in time. For spatial PPs, such as, say, *above the cupboard*, *in front of the cupboard* or *in the cupboard*, it has been suggested [Bierwisch, 1983] that they refer to a certain *spatial region* and locate the relevant entity as lying within this region. The region is determined as one which stands to the referent of the preposition-governed NP in a relation expressed by the given preposition. For example, in the case of *above the cupboard* the region consists of portions of space which are encountered when one moves vertically upwards from any part of the top of the cupboard. (In normal situations, where the cupboard is indoors, it is the space between the cupboard and the ceiling.) Similarly for *in front of the cupboard*, *in the cupboard* and so on.

On this view the adjunction of a spatial location PP involves two relations linking the referent of the NP it governs to the entity which it serves to locate (i) the relation expressed by the preposition, which holds between the referent of the NP and the region denoted by the PP as a whole, and (ii) the relation of spatial inclusion between that region and the entity that is being located; this second relation is the semantic correlate of the syntactic relation between adjunct and adjunction site.[26]

Not only do we endorse this proposal about the interpretation of spatial PPs, we also propose that the same analysis be adopted for PPs which express temporal location. According to the extended proposal the PP *after the exam* denotes a certain region of time — an interval which extends from the exam into the future, with an intrinsically vague upper bound — and when the PP is used in the way it is in (91), i.e. as locating predicate of the described event, it imposes a locating constraint on this event via the condition that the event's location time is included in the temporal region denoted by the PP.

Instead of presenting the DRS for the sentence *After the exam Fred bought a lawn mower.* itself we give, in order to catch two birds with a single DRS in (92.b) the representation of a sentence in which this PP serves to locate a state rather than an event. This sentence is given in (92.a) — the VP *have a headache* is generally assumed to have stative aspect.

(93) a. After the exam Fred had a headache.

b.

$$\boxed{\begin{array}{c} \mathrm{n\ f\ s\ t}_{loc}\ \mathrm{t}_{ex}\ \mathrm{t}_{reg} \\ \mathrm{Fred(f)} \\ \text{“the exam”}(\mathrm{t}_{ex}) \\ \mathrm{t}_{loc} \prec \mathrm{n} \\ \mathrm{t}_{ex} \supset\!\subset \mathrm{t}_{reg} \\ \mathrm{t}_{loc} \subseteq \mathrm{t}_{reg} \\ \mathrm{t}_{loc} \subseteq \mathrm{s} \\ \mathrm{s}: \text{“have-a-headache”}(\mathrm{f}) \end{array}}$$

The present treatment of the semantics of temporal adverbials has one consequence which deserves special mention in view of the amount of attention which this matter has received in the literature. To the sentence *Yesterday Fred had a headache.* our treatment assigns truth conditions according to which the state of Fred having a headache is merely required to overlap with yesterday. Thus the analysis does not require that the headache lasted all day. For the present example this seems to be all to the good, but we hasten to add that the issue is more involved than this one example reveals.

We conclude this discussion of the role of temporal adverbials with four remarks of a more general methodological nature.

[26]Note that for a PP whose preposition is *in* (such as *in the cupboard*) this analysis has a semblance of redundance, since inclusion is expressed by the preposition as well as by the syntactic adjunction configuration. But of course this incidental duplication does not speak against the proposal as such.

1. As the interaction between times and adverbs has been analysed here, it involves a combination of several constraints on one and the same entity (viz. the location time represented by the discourse referent t_{loc}), with one constraint contributed by the adverb and the other by tense. This means in essence that the mechanism of tense-adverb interaction involves a form of semantic unification. In fact, it was because of its unification-like character that the analysis of this interaction has had a decisive influence on DRT's general conception of the syntax-semantics interface.

2. A typical feature of unification is that it can fail when the constraints that need to be unified are incompatible with each other. Interactions between tenses and temporal adverbs manifest this typical feature of unification-based processes too. An example is the sentence in (94).

(94) Yesterday Fred will buy a lawn mower.

Here the constraint imposed on t_{loc} by *yesterday* requires it to precede n, while the constraint imposed by the future tense forces it to follow n. Consequently interpretation aborts, with the effect that the sentence is felt to suffer from a special kind of "semantic ungrammaticallity".

3. So far we have considered a couple of examples (those in (85)) which demonstrated the dynamic poperties of tense (and especially its backwards dynamic, or "anaphoric" properties) and after that a number of examples where temporal location is constrained by a clause-internal adverb (and where the constraints imposed by adverb and tense have to be consistent). What happens in situations where both those mechanisms are applicable? The unification perspective would suggest that the same consistency constraints apply in these cases too: if there is a conflict between the constraints imposed by the adverb and the relation in which the location time stands to the context-supplied reference time, then the sentence tends to be uninterpretable or at least to be judged infelicitous. One type of example of this are sentences in which the adverb denotes a time which is located well before the contextual reference time and where the tense is a simple past. Such sentences often sound bad, or seem incomprehensible. (In such cases the past perfect is usually required, or at least it is preferred over the simple past. The reasons for this will become clear when we discuss our next example.)

But although the constraints contributed by reference time and adverb often seem to lead in such cases to conflicts which render the sentence infelicitous, there is nevertheless an asymmetry between them. Adverbial constraints tend to overrule contextual contraints. This is no surprise given that the principles which govern adverbial reference are much more clearly defined (and therefore less amenable to reinterpretation on the spur of the moment) than those which govern the links beween the new sentence and its context. The upshot of this is that to the extent that the system of temporal location we find in a language like English can be regarded as unification-based, the unification involved is one that allows for constraint prioritisation. In other words, we are dealing with a form of default unification (cf. [Briscoe *et al.*, 1993; Lascarides and Copestake, 1999].)

4. The interaction between tense and temporal adverb has also been of central importance for the overall structure of the DRT-account of temporal reference that is the particular topic in the present section. As we noted at the beginning of this section, one of the reasons why temporal logics do not provide a satisfactory framework for the analysis of temporal reference in natural language is their lack of any devices for explicit reference to times. We cited temporal adverbs as salient examples among the expressions of natural languages for which it is obvious that they do explicitly refer to times. The way in which adverbs and tenses work together in locating eventualities along the time axis is important in this connection insofar as it indicates that treating tenses and temporal adverbs separately, using one representational framework to deal with tense and another to deal with adverbs, would be a hopeless undertaking. We need representations which contain terms standing for times to represent the contribution of the tenses no less than we need such terms to represent the contributions that are made by the adverbs.

Perspectival Shift and the Two-Dimensional Theory of Tense

So far we have considered a couple of tenses which can be analysed without reference to temporal perspective points. (Other tenses which allow for a similar analysis are the present tense and simple future tense of English (recall (86)) and similarly the présent and the futur of French.) But this is not true in general. That there are tenses which require a more complicated analysis is arguably the most salient feature of Reichenbach's theory of tense. Reichenbach showed that when a sentence in the simple past is followed by a sentence in the past perfect, the eventuality described by the latter is typically understood as preceding the former: the first sentence provides a "past reference time" for the interpretation of the past perfect of the second sentence and the past prefect locates its eventuality in the past of this past perspective time. The following example illustrates this principle. At the same time it shows why it is necessary to distinguish between temporal perspective time and reference time.

(95) Luigi was writing to the Department Chairman. He had applied for the job without much hope. But the Committee had invited him for a talk, he had given a perfect presentation, they had offered him the job

and he had accepted. Now he was worried about what he was going to teach.

(95) begins with a sentence S_1 in the past progressive. S_1 is followed by a sequence of five sentences S_2, S_3, S_4, S_5, S_6 in the past perfect. The passage ends with a sentence S_7 in the simple past.

The first aspect of (95) that matters here is the interpretation of the past perfects in S_2-S_6. We start by looking at S_2. In the context provided by the sentence S_1 which precedes it S_2 is understood as describing an event e_2 situated in the past of the location time t_1 of the eventuality s_1 described by S_1. The observation that this is so, we just noted, was the central insight which led Reichenbach to his "two-dimensional" theory. In the present account, Reichenbach's analysis of the past perfect has been taken over, except that here it is the temporal perspective time which plays the intermediate role between event time (i.e. our location time) and utterance time. This role cannot be played by what we have called the temporal reference time, as the reference time may be needed in a different capacity. To see this consider the second past perfect sentence of (95), i.e. S_3. The interpretation of this sentence involves temporal location of the described event e_3 in the past of t_1, and we may assume that the same mechanism is responsible for this that also locates e_2 before t_1. On the other hand e_3 is understood as following e_2, and the mechanism responsible for this is strongly reminiscent of what we saw in our discussion of (85.b). There the second of a pair of sentences in the simple past was interpreted as describing an event whose location time stood in a relation ρ of temporal succession to the location time t_1 of the event described by the first sentence, and in the interpretation of the tense of the second sentence t_1 was assumed to play the role of temporal reference time. We claim that a similar relationship holds between the location times of the events introduced by the second and third sentence of (95), and that like in the case of (85.b) the location time of S_2 acts as reference time in the interpretation of S_3.

According to this analysis the interpretation of S_3 involves both a temporal reference time and a perspective time. Since the new location time t_3 is assumed to stand to reference time and perspective time in distinct relations — the relation to the reference time we have assumed is that of the reference time preceding t_3, whereas t_3 precedes the perspective time, as it does for any past perfect — reference time and perspective time must be distinct. Hence the need for two notions rather than one.

Note that for each of the sentences S_3, ..., S_6 reference time and perspective time are distinct. Moreover, while the perspective time remains constant, the reference time changes from sentence to sentence. This is a typical feature of extended flashbacks. These remarks evidently do not solve the problem how perspective times are chosen in general. For one thing, not every sequence of sentences in the past perfect following a sentence in the simple past constitutes a single extended flashback. Sometimes we find flashbacks within flashbacks, and in such cases the perspective time for the sentence or sentences of the embedded flashback is not the location time of the last simple past sentence but that of an earlier past perfect sentence. However, the question when we are dealing with a single flashback and when with an embedding of one flashback within another once again depends on factors on which the account we have sketched has no purchase. Thus the choice of perspective time is (for the reason given as well as others) a problem that our account can deal with only to a first approximation — just as we found this to be the case for the specification of the relation ρ.

In (96) and (97) we present the relevant stages in the interpretation of S_3. The DRS in (96) gives complete presentations for S_1 and S_2 together with a preliminary representation for S_3 which displays the two presuppositions triggered by the past perfect. One of these concerns the relation to the reference time and is identical with the presupposition of (87), while the other has to do with the relation to the perspective time.[27]

$$(96)\quad \begin{array}{|c|}\hline \mathrm{n\ l\ c\ t_1\ s_1\ j\ t_2\ e_2} \\ \text{Luigi(l)} \\ \text{“the Department Chairman”(c)} \\ \mathrm{t_1 \prec n} \\ \mathrm{t_1 \subseteq s_1} \\ \mathrm{s_1}\text{: PROG(}{}^\wedge\text{e. e: write-to(l,c))} \\ \text{“the job”(j)} \\ \mathrm{t_2 \prec t_1} \\ \mathrm{e_2 \subseteq t_2} \\ \mathrm{e_2}\text{: apply-for(l,j)} \\ \text{“without-much-hope”}(\mathrm{e_2}) \\ \hline \end{array}$$

$$\left\langle \left\{ \begin{array}{|c|}\hline \mathrm{r_3} \\ \rho_3(\mathrm{r_3,t_3}) \\ \hline \end{array}, \begin{array}{|c|}\hline \mathrm{p_3} \\ \mathrm{p_3 \prec n} \\ \hline \end{array} \right\}, \begin{array}{|c|}\hline \mathrm{C\ t_3\ e_3} \\ \text{“the Committee”(C)} \\ \mathrm{t_3 \prec p_3} \\ \mathrm{e_3 \subseteq t_3} \\ \mathrm{e_3}\text{:”invite-for-a-talk”(C,l)} \\ \hline \end{array} \right\rangle$$

To obtain the final representation of the first three sentences of (95) the two presuppositions of (96) must still be resolved. How they should be resolved has already been stated: r_3 must be identified with t_2 and p_3 with t_1, while

[27]N.B. In (96) we have simplified the representation of anaphoric pronouns (such as the *he* of S_2) by substituting the discourse referents for their anaphoric antecedents into their argument positions (instead of introducing a distinct discourse referent for the pronoun together with an equation which enforces coreference between it and the antecedent discourse referent). This too is a practice which from now on we will adopt whenever it suits us.

the relation ρ is to be specified as "$\prec$", so that the condition "$\rho_3(r_3,t_3)$" turns to "$t_2 \prec t_3$". These resolutions lead to the representation in (97).

(97)

n l c t_1 s_1 j t_2 e_2 C t_3 e_3
Luigi(l)
"the Department Chairman"(c)
$t_1 \prec n$
$t_1 \subseteq s_1$
s_1: PROG(^e. e: write-to(l,c))
"the job"(j)
$t_2 \prec t_1$
$e_2 \subseteq t_2$
e_2: apply-for(l,j)
"without-much-hope"(e_2)
"the Committee"(C)
$t_1 \prec n$
$t_3 \prec t_1$
$t_2 \prec t_3$
$e_3 \subseteq t_3$
e_3:"invite-for-a-talk"(C,l)

The interpretation of the sentences S_4-S_6 proceeds in the same way as that of S_3 and requires no further comment. But the last sentence S_7 of (95) presents a new problem, which is connected with the occurrence in it of the word *now*. This is a problem of a kind which we have not yet encountered and which merits separate discussion.

Apart from the question raised by *now*, the representation of S_7 also presents some difficulties which are orthogonal to the concerns of this section. These have to do with the embedded question *what he was going to teach*. We finesse them by considering instead of S_7 the simpler sentence

(98) Now he was worried.

We will refer to this sentence as S_7'. (So the revised version of (95) consists of the sentences S_1,S_2,S_3,S_4,S_5,S_6,S_7'.) Looking at S_7' rather than S_7 allows us to focus attention on the issue that matters in the present context.

The problem arises from the fact that in S_7' *now* occurs in the presence of the past tense. It has been claimed that *now* is an indexical adverb which always refers to the time at which it is uttered [Kamp, 1971]. If this were true without qualification, then the interpretation of S_7' should abort, since the constraints imposed on its location time t_7 by tense and adverb would be incompatible. The fact that in the given context S_7' is not uninterpretable indicates that (at least in this type of context) *now* can be used to refer to times which lie in the past of the utterance time.

An inspection of those cases where *now* can refer to a time other than the one at which it is actually uttered suggests that it is referring in such cases to a time that can be regarded as a kind of "displaced utterance time" (or, in slightly different terms, to the time of a "view point", or "perspective" which has been shifted away from the real utterance time).[28]

The appeal to "shifted view points" may have a plausible ring to it, but as it stands it is too vague to serve in a formal theory of temporal interpretation of the kind we are pursuing. So we are facing the question: How can this notion be made more concrete? [Kamp and Reyle, 1993], following an earlier proposal in the unpublished manuscript [Kamp and Rohrer, 1983b], proposes that the perspectival shifts that are involved in the reference of *now* to a time in the past of n are the same as those involved in the interpretation of a past perfect sentence (viz. as placing its eventuality in the past of a time that itself is in the past of the utterance time) and thus that the same notion of perspective time we have just been invoking for the interpretation of S_2-S_6 is also the one to be invoked in the interpretation of S'_7. According to this proposal the interpretation of *now* involves two possibilities, as stated in (99).

(99) *now* either refers to the utterance time, or else to a time which plays the role of perspective time in the interpretation of the sentence to which it belongs.

Given this assumption it follows that the interpretation of S'_7 follows a pattern that closely resembles that of S_2-S_6. Again the interpretation requires the choice of a perspective time and once again the intuitively right candidate for this is the location time t_1 introduced in the interpretation of S_1. We assume therefore that the preliminary representation of S'_7 involves the same pair of presuppositional components that are also part of the preliminary representations of S_2-S_6. This representation is given in (100).

$$(100)\ \left\langle \left\{ \boxed{\begin{array}{c} r_7 \\ \rho_7(r_7,t_7) \end{array}}, \boxed{\begin{array}{c} p_7 \\ p_7 \prec n \end{array}} \right\}, \boxed{\begin{array}{c} t_7\ e_7\ t_{now} \\ t_7 \subseteq s_7 \\ r_7 \subseteq t_7 \\ t_{now} = p_7 \\ t_7 \subseteq t_{now} \\ s_7\text{:"be-worried"}(l) \end{array}} \right\rangle$$

Resolution of the perspective time p_7 should, we said, again be to t_1. What about the reference time r_7? Before we try to answer this question let us see what we know about the location time t_7 of the state which S_7

[28]For reasons which we make no effort to explain here such shifts seem to occur almost exclusively in the direction of the past; cases where *now* refers to some time in the future of the utteramce time appear to be marginal. But see [Sandström, 1993].

describes. Identifying p_7 with t_1 means that *now* is construed as referring to t_1. At the same time *now*, as a temporal adverb, serves as a constraint on the new location time t_7; thus $t_7 \subseteq t_1$. So all that we need and can hope for from the resolution of the presupposition concerning r_7 is that the result is consistent with this interpretation. As before, resolution of this presupposition involves (i) specifying ρ_7, and (ii) finding an antecedent for r_7. As regards (i) note that tense and aspect of S'_7 are like those of the second sentence of (85.a): the sentence has stative aspect and its tense is in the simple past. This suggests that once again the relation should be inclusion: $r_7 \subseteq t_7$. This leaves r_7. It is clear that the only possible antecedent for r_7 within the context provided by S_1-S_6 which is consistent with the constraints that have been established already is the location time t_1. Resolving r_7 accordingly means that this time reference time and perspective time both get identified with t_1 and thus that they coincide. (101) gives (in abridged form) the representation for (95) that results from these resolutions of the two presuppositions of (100).

The choice of t_1 as reference time for S'_7 indicates that the determination of reference times is in general more complex than was revealed by the examples that have so far been discussed in this section. Choosing the location time t_1 of the first sentence S_1 after the sentences S_2, ..., S_6 have intervened reflects the perception that the flashback S_2, ..., S_6 has come to an end and that S'_7 returns to the point of the story which had been reached with S_1 and then interrupted. Here we see a correspondence between choice of reference time and narrative structure, a correspondence which once more transcends the scope of the account as it has been presented.

Discussion: Resolution of Reference Time, Perspective Time and their Relations; Perspective Time and Perspectival Shifts; A General 2-Dimensional Theory of Tense

The treatment of (95) which we have presented here leaves some questions unanswered and suggests some new ones of its own.

1. First, as we have stated it our account of (95) contains a number of loose ends. The most serious of these concern (i) the principles governing the identification of reference time r and perspective time p; and (ii) the specification of the relations in which r and p stand to location time and utterance time. As we noted in the discussion after our analysis of (85), a central problem, and one on which the present account has nothing to say, is that the specification of ρ often depends on other, "pragmatic" factors besides those we have considered. In our discussion of the interpretation of S_3-S_6 of (95) we observed that the same applies to the choice of perspective time and we concluded in connection with our discussion of S'_7 that the choice of reference time gives rise to similar problems. The same proves to

(101)

$$
\begin{array}{|c|}
\hline
\mathrm{n\ l\ c\ t_1\ s_1\ j\ t_2\ e_2\ C\ t_3\ e_3\ ...\ t_5\ t_6\ e_6\ t_7\ e_7\ t}_{now} \\
\hline
\text{Luigi(l)} \\
\text{“the Department Chairman”(c)} \\
\mathrm{t_1 \prec n} \\
\mathrm{t_1 \subseteq s_1} \\
\mathrm{s_1}\text{: PROG(}^\wedge\text{e. e: write-to(l,c))} \\
\text{“the job”(j)} \\
\mathrm{t_2 \prec t_1} \\
\mathrm{e_2 \subseteq t_2} \\
\mathrm{e_2}\text{: apply-for(l,j)} \\
\text{“without-much-hope”(}\mathrm{e_2}\text{)} \\
\text{“the Committee”(C)} \\
\mathrm{t_1 \prec n} \\
\mathrm{t_3 \prec t_1} \\
\mathrm{t_2 \prec t_3} \\
\mathrm{e_3 \subseteq t_3} \\
\mathrm{e_3}\text{:”invite-for-a-talk”(C,l)} \\
\vdots \\
\mathrm{t_6 \prec t_1} \\
\mathrm{t_5 \prec t_6} \\
\mathrm{e_6 \subseteq t_6} \\
\mathrm{e_6}\text{: accept(l,j)} \\
\\
\mathrm{t_7 \subseteq t_1} \\
\mathrm{t_7 \prec n} \\
\mathrm{t}_{now} = \mathrm{t_1} \\
\mathrm{t_7 \subseteq s_7} \\
\mathrm{s_7}\text{:”be-worried”(l)} \\
\hline
\end{array}
$$

be true with regard to the resolution of r and p. It appears that if we want to make substantial further progress on these problems we need a framework in which these other factors can be treated in a systematic way. As it stands DRT does not provide this framework.

2. Another question which naturally arises in the context of what we have said about (95) concerns the need for the notion of perspective time. We argued in relation to the non-initial sentences of extended flashbacks — in the case of (95): sentences S_3, ..., S_6 — that their interpretation involves linking to two different times from the context. Since the reference time cannot be responsible for both these links at once, we said, a further notion is needed. But does that really follow? There might be an alternative way of dealing with this problem, viz by maintaining that the past perfect (and possibly other tenses as well) requisitions the reference time for its own needs, and thereby creates the possibility of choosing a further, "secondary" reference time, which can take over the task that is accomplished by the "primary" reference time in cases where the tense does not come with such special needs (e.g. when it is a simple past). On the face of it this may seem to be nothing more than a superficial variant of the account we have presented. But it is connected with a more substantive issue. Even if we adopt this variant there is still a need for some notion of perspective time in connection with "pseudo-indexical" uses of indexical adverbs, like that of *now* in the last sentence of (95). But shifted references of *now* and run-of-the-mill past perfects would no longer be represented as involving one and the same conceptual operation (that of choosing a past perspective point). And this is a disentanglement that some would welcome. It may be added in this connection that not all cases of shifted reference by indexicals appear to be of the same kind. For instance, there are subtle differences between the kind of perspectival shift we find with a word like *now* and the shifts involved in shifted reference of adverbs like *yesterday* or *tomorrow*. (See [Kamp and Rohrer, 1983b]). How many different notions of perspective will be needed eventually to do justice to these differences remains open.

3. The examples we have discussed in detail involved only two tense forms, the simple past and the past perfect. (The past progressive, we said, should be analysed as a combination of the simple past and an aspectual operator which transforms a verb into its progressive form.) And for only one tense form, viz., the past perfect, did our analysis require perspective time. It is a natural question for which other tenses (if any) perspective times are needed as well. Answers to this question lie somewhere between a lower and an upper bound. The lower bound consists of a small set of tenses which includes besides the past perfect also the the "future of the past", as we find it in the second sentence of (102).

(102) On the 3rd Powell arrived in Brussels. On the 4th he would be in London and on the 5th in Berlin.

(With such future-of-the-past sentences the location time of the described eventuality follows the perspective time while the perspective times precedes n.) The upper bound is the set consisting of all tense forms. A proposal to the effect that perspective time is involved in the analysis of all tenses can be found in [Kamp and Reyle, 1993] (and in the unpublished [Kamp and Rohrer, 1983b] for the tenses of French). Tenses which in a lower bound account would be treated as not involving perspecive time (as one assumed for the simple past in the analyses given here) are analysed in this proposal as locating the perspective time at the utterance time. Since the proposal uses perspective time both to account for the tenses and for perspectival reference shifts for words like *now* (just as was assumed above), a consequence is that the simple past tense is ambiguous between an analysis where the perspective time coincides with n — see the treatment of (85.a,b) above — and one where the pespective time lies in the past of n and coincides with the location time — see the above treatment of sentence S'_7. This consequence has been perceived as undesirable and seen as a further argument against assuming that perspective times serve in this dual capacity. A different explanation of the possibility of shifted reference of words like *now* can be found in [Roßdeutscher, 2000].

The questions raised under points 1.-3. above are of prime importance for linguistics. But they carry no implications for the form of the DRS-language that is needed to represent temporal information. As can be seen from the examples we have presented, all reference to r and p has disappeared from the representation when the DRS for a sentence or discourse has reached its final form. Since the form of these final DRSs — i.e. of the "formulas" of our DRS language — is independent of the details of DRS construction in which the alternative accounts alluded to in the discussion above differ from the account we have presented, these details will be of lesser interest to those readers who are primarily concerned with form and meaning of the final representation language.

Temporal Quantification

So far we have looked at the interaction between tenses and temporal "locating" adverbs. These adverbs, we argued, denote certain periods (or "regions") of time, within which the location time of the described eventuality is situated. But not all temporal adverbs function this way. Just as among NPs we find besides the definite and indefinite ones, which have some sort of referential status, also quantificational NPs, so we find quantificational temporal adverbials besides those locating adverbs which contribute to the interpretation of the sentences in which they occur just one particular time. Quantificational temporal adverbials come in two main forms: (i) prepositional phrases whose NP is quantificational, such as *on every Sunday*, *on most Sundays between June 15th and August 31st*, *after many parties thrown by Mary*, etc.; and (ii) quantifying adverbs such as *often*, *always*, *usually*, *regularly*.

The question how sentences containing quantifying temporal adverbials should be represented is somewhat easier for adverbials of type (i). What we would expect in this case is that the representation of a sentence with a quantifying temporal PP stands to that of a corresponding sentence in which the quantifying NP of the PP has been replaced by a referential NP in the same relation that, say, the representation of a sentence with a quantificational subject stands to that of the sentence we get by replacing this subject NP by a referential one. Compare for instance the following four sentences.

(103) a. The Dream of Gerontius is boring.

b. Every choral work of Elgar is boring

c. On Sunday Mary went to see her aunt.

d. On every Sunday between June 15-th and August 31-st, 2001 Mary went to see her aunt.

(104.a) and (104.b) give DRSs for (103.a,b) in accordance with the proposals of Sections 2 and 3.1 and (104.c) gives a representation of (103.c) according to the proposals that have already been made in the present subsection:

(104) a. $\boxed{\begin{array}{c} \text{d} \\ \text{“The-Dream-of Gerontius”}\,(\text{d}) \\ \text{boring}(\text{d}) \end{array}}$

b. $\boxed{\begin{array}{c} \text{e} \\ \text{Elgar}(\text{e}) \\ \boxed{\begin{array}{c} \text{x} \\ \text{“choral-work-of”}\,(\text{x,e}) \end{array}} \; \left\langle \begin{array}{c} \forall \\ \text{x} \end{array} \right\rangle \; \boxed{\text{boring}(\text{x})} \end{array}}$

c.
$$\begin{array}{|c|}\hline \mathrm{n\ m\ a\ t}_{loc}\ \mathrm{e\ t}_s\ \mathrm{t}_{adv} \\ \text{Mary(m)} \\ \text{aunt(a,m)} \\ \text{“Sunday”}(\mathrm{t}_s) \\ \mathrm{on}(\mathrm{t}_{adv},\mathrm{t}_s) \\ \mathrm{t}_{loc} \prec \mathrm{n} \\ \mathrm{t}_{loc} \subseteq \mathrm{t}_{adv} \\ \mathrm{e} \subseteq \mathrm{t}_{loc} \\ \text{e: “go-to-see”(m,a)} \\ \hline \end{array}$$

If the representation for (103.d) is to be related to (104.c) in the way that (104.b) stands to (104.a), it should be something like the one given in (105).

(105)
$$\begin{array}{|c|}\hline \mathrm{n\ m\ a} \\ \text{Mary(m)} \\ \text{aunt(a,m)} \\ \begin{array}{|c|}\hline \mathrm{t}_s \\ \text{“Sunday-between...”}(\mathrm{t}_s) \\ \hline \end{array} \;\left\langle \begin{array}{c} \forall \\ \mathrm{t}_s \end{array} \right\rangle\; \begin{array}{|c|}\hline \mathrm{t}_{loc}\ \mathrm{e\ t}_{adv} \\ \mathrm{on}(\mathrm{t}_{adv},\mathrm{t}_s) \\ \mathrm{t}_{loc} \prec \mathrm{n} \\ \mathrm{t}_{loc} \subseteq \mathrm{t}_{adv} \\ \mathrm{e} \subseteq \mathrm{t}_{loc} \\ \text{e: “go-to-see”(m,a)} \\ \hline \end{array} \\ \hline \end{array}$$

(105) is adequate insofar as it captures the truth conditions of (103.d) correctly. But it provides no insight into the question which has been high on our agenda so far: how do tense and temporal adverb interact to produce such interpretations?

Intuitively it seems clear that the tense of (103.d) is relevant to the interpretation of the sentence insofar as it locates the possible values of the sentence-internally bound variable t_{loc} in the past of n. In other words, whenever the quantificationally bound variable t_s takes a value satisfying the restrictor predicate *Sunday between 15-06-2001 and 31-08-2001*, and t_{loc} is a time included within the time t_{adv} (which in this case will coincide with the value of t_s), then t_{loc} must precede n. This is consistent with speakers' intuitions about use and meaning of (103.d): if we assume that (103.d) is uttered at a date after 31-08-2001 (such as, say, March 2003), in which case all values for t_{loc} which satisfy the restrictor predicate also satisfy the constraint imposed by the past tense (i.e. are in the past of n), then the sentence is used felicitously. But when the sentence is uttered at some time before this date, then some values for t_{loc} will not lie before n; and, indeed, such an utterance would be perceived as incoherent or strange. By the same token, (106) would be incoherent at any time when (103.d) can be used coherently.

(106) On every Sunday between June 15^{th} and August 31^{st}, 2001 Mary will go to see her aunt.

In the light of these observations, together with what has been said about the interaction between tenses and referential temporal adverbs earlier, the following would appear to be a natural hypothesis about the way in which the tenses of sentences like (103.d) and (106) and the adverbs of these sentences interact:

(107) The tense contributes its constraints to the nuclear scope of the duplex condition introduced by the quantificational adverb.

In fact, we already used this hypothesis in the construction of the DRS in (105), where the condition "$\mathrm{t}_{loc} \prec \mathrm{n}$", contributed by the simple past of (103.d), is one of the conditions in the nuclear scope DRS.

Unfortunately, however, (107) isn't correct in general. It fails for sentences in the present tense. Consider (108).

(108) This week the patient is checked every half hour.

Let us assume that (108) is uttered on a Wednesday. It then asserts that at half hourly intervals throughout the week to which the given Wednesday belongs there are occurrences of events of the described type (i.e. of the patient being checked). Some of these events are situated in the past of the utterance time, some of them in the future of it, and perhaps one is going on at the very moment when the statement is made. Though we haven't discussed the constraints imposed by the English present tense explicitly, we trust that the reader is prepared to accept this much: not all these different temporal relations in which patient-checking events stand to the utterance time n are compatible with the constraints it imposes. (This follows in particular if we assume that the contribution of the present tense is as given in (86). In actual fact the English present tense covers a somewhat wider set of possibilities than (86) allows for, but the present point is not affected.) If not all the events of which (108) asserts that they took, take or will take place satisfy the constraint which the present tense imposes, then (107) is refuted. What then is the way in which tense and quantificational temporal adverbs interact? And in particular, how can we explain that (108) is an acceptable sentence? The answer we propose is the following:

(109) In quantificational statements like those in (103.d), (106) or (108) the tense of the sentence locates the temporal interval within which the times from the domain of the quantification are included.

In the case of (103.d) this interval is located entirely in the past of the utterance time, whence a past tense is appropriate there. Likewise, with future tense substituting for past tense, for (106). In the case of (108) the interval straddles the utterance time, and this requires the present tense.

To find a justification for (109) we may look in either or both of two directions. The first involves the assumption that duplex conditions can function as characterisations of eventualities. We refer to such eventualities as “quantificational states”. On this assumption the duplex condition in (105) can be construed as the description of a state s as shown in (110):

$$\text{(110) s:} \quad \boxed{\begin{array}{c} t_s \\ \text{“Sunday-...”}(t_s) \end{array}} \left\langle \begin{array}{c} \forall \\ t_s \end{array} \right\rangle \boxed{\begin{array}{c} t_{loc}\ e\ t_{adv} \\ \hline on(t_{adv},t_s) \\ t_{loc} \prec n \\ t_{loc} \subseteq t_{adv} \\ e \subseteq t_{loc} \\ e\text{: “go-to-see”}(m,a) \end{array}}$$

The full representation of (103.d) of which (110) is part includes in addition the introduction of s as a member of its DRS universe. Moreover, just as any other eventuality, s is assumed to come with its own location time t, and it is this location time that is assumed to be the time that is constrained by tense. With these assumptions (105) turns into (111):

$$\text{(111)} \quad \boxed{\begin{array}{c} n\ m\ a\ t\ s \\ \hline Mary(m) \\ aunt(a,m) \\ t \prec n \\ t \subseteq s \\ \text{s:} \quad \boxed{\begin{array}{c} t_s \\ \text{“Sunday-...”}(t_s) \end{array}} \left\langle \begin{array}{c} \forall \\ t_s \end{array} \right\rangle \boxed{\begin{array}{c} t_{loc}\ e\ t_{adv} \\ \hline on(t_{adv},t_s) \\ t_{loc} \subseteq t_{adv} \\ e \subseteq t_{loc} \\ e\text{: “go-to-see”}(m,a) \end{array}} \end{array}}$$

Like (105), (111) correctly captures the truth conditions of (103.d). But we aren’t out of the woods yet. This becomes clear when we consider sentences in which the quantificational temporal adverb is in the scope of another temporal adverb, as it is for instance in (108). In (108) the time specified by the “outer” adverb *this week* functions as an additional restriction on the quantification expressed by the “inner” adverb: we are talking about events one half hour apart throughout the week containing the utterance time.

In order to keep the connection with the representations of adverbial quantifications we have considered so far (i.e. (105) and (111)) as transparent as possible, let us look, not at (108), but at the following variant of (103.d):

(112) Last summer Mary went to see her aunt (on) every Sunday.

Suppose we try to construct a representation for (112) along the lines of (111). The additional matter we now have to deal with is the adverb *last summer*. In the light of what we have said above about how referential temporal adverbs contribute their semantics, *last summer* should constrain the relevant location time as included within the period t which the adverb denotes. (Somewhat simplified, if t_{ls} is the summer of the year preceding the one in which the utterance is made, and t is the relevant location time, then the constraint contributed by the adverb should be the condition "$t \subseteq t_{ls}$".)

What is the relevant location time in this case? It is easy to see that it cannot be the one which in (110) and (111) appears within the nuclear scope of the duplex condition. For that would clearly lead to the wrong truth conditions. The only other possibility is that the relevant location time is the location time t of the quantificational state s. However, as it stands the condition "$t \subseteq t_{ls}$", in which t is this location time, does not give us what we want either. For the only conclusion which it allows us to draw is that s overlaps with the denotation of *last summer*. And that is too weak. What we want is this: the temporal quantification is restricted to the period denoted by *last summer*.

To get this stronger implication we need a pair of further stipulations:

(113) (i). The duration of a quantificational state coincides both with its location time and its adverb time; moreover,

(ii). the quantification which characterises a quantificational state is by definition restricted to the state's duration.

Given (113) we get the following "upgraded" representation for (112):

(114)

n m a t s t_{ls}

Mary(m)
aunt(a,m)
$t \prec n$
$t = \text{dur}(s)$
"last summer"(t_{ls})
$t_{ls} = \text{dur}(s)$

s: [t_s | "Sunday"(t_s), $t_s \subseteq t$] $\langle \forall\ t_s \rangle$ [t_{loc} e t_{adv} | on(t_{adv},t_s), $t_{loc} \subseteq t_{adv}$, $e \subseteq t_{loc}$, e: "go-to-see"(m,a)]

But what is the justification for the assumptions made in (113)?

We can get closer to such a justification by following the second one of the two directions hinted at. This direction has to do with the contextual constraints that quantification has been observed to be subject to in general. We noted in Section (3.3) that quantification often involves tacid restrictions and we followed the proposal of [von Fintel, 1994] and others to represent these in the form of an additional restriction on the bound variable of the quantification, involving an initially unspecified predicate C. C is introduced as part of a presuppostion which requires resolution in the light of contextual information.

When the variable bound by the quantifier ranges over times, the resolution of C often takes on a special form: that of a "frame interval" within which the values of the bound variable are temporally included. (In such cases resolution of C may involve other factors as well, a point to which we turn below.) Moreover, when the quantificational temporal adverbial is within the scope of another temporal adverb — as it is in (108) or (112) — it is the outer adverb which specifies the frame interval for the quantification expressed by the inner adverb. (In such cases an expression belonging to the sentence itself accomplishes what in its absence would be the task of the context. Recall what was said on this score in the section on tenses and locating adverbs.)

In this way the constraint contributed by the outer temporal adverb becomes part of the restrictor of the quantification, which is where it is wanted. (115.a) gives the representation of (112) before resolution of the restrictor predicate C and (115.b) the result of resolving C to the referent of *last summer*.

Like (114), (115.b) renders the truth conditions of (112) correctly. But once more we need to ask: what could be the deeper justification for the assumptions on which the new representation rests? That the outer adverb can serve as a source for the specification of C seems plausible enough. But even if we asume that it can serve this purpose, that is not the same thing as showing that it *must* be understood in this capacity. Perhaps it could be argued that this is the only meaningful function that the outer adverb could have in a sentence like (112), so that the necessity of its contribution to the restrictor of the adverbial quantification becomes an instance of "full interpretation": each potentially meaningful constituent of a sentence must make a meaningful contribution to the whole. But it is unclear to us how this intuitive principle could be made more precise.

The point we have reached can be summarised as follows. We have looked at two mechanisms which could be held responsible for the interaction between quantificational temporal PPs, tenses and other temporal adverbs: (i) "reifying" the quantifications expressed by quantifying temporal PPs as "quantificational states" whose types are given by the duplex conditions

(115)

a.

$$\begin{array}{|c|}\hline \mathrm{n\ m\ a\ t\ s\ t}_{ls} \\ \text{Mary(m)} \\ \text{aunt(a,m)} \\ \mathrm{t} \prec \mathrm{n} \\ \mathrm{t = dur(s)} \\ \text{“last summer”}(\mathrm{t}_{ls}) \\ \mathrm{t}_{ls} = \mathrm{dur(s)} \\ \left\langle \left\{ \begin{array}{|c|}\hline \mathrm{C} \\ \mathrm{P(C)} \\ \hline \end{array} \right\},\ \mathrm{s}:\ \begin{array}{|c|}\hline \mathrm{t}_s \\ \text{“Sunday”}(\mathrm{t}_s) \\ \mathrm{C}(\mathrm{t}_s) \\ \hline \end{array} \ \left\langle \begin{array}{c} \forall \\ \mathrm{t}_s \end{array} \right\rangle \ \begin{array}{|c|}\hline \mathrm{t}_{loc}\ \mathrm{e}\ \mathrm{t}_{adv} \\ \mathrm{on}(\mathrm{t}_{adv},\mathrm{t}_s) \\ \mathrm{t}_{loc} \subseteq \mathrm{t}_{adv} \\ \mathrm{e} \subseteq \mathrm{t}_{loc} \\ \mathrm{e}:\ \text{“go-to-see”(m,a)} \\ \hline \end{array} \right\rangle \\ \hline \end{array}$$

b.

$$\begin{array}{|c|}\hline \mathrm{n\ m\ a\ t\ s\ t}_{ls} \\ \text{Mary(m)} \\ \text{aunt(a,m)} \\ \mathrm{t} \prec \mathrm{n} \\ \mathrm{t = dur(s)} \\ \text{“last summer”}(\mathrm{t}_{ls}) \\ \mathrm{t}_{ls} = \mathrm{dur(s)} \\ \mathrm{s}:\ \begin{array}{|c|}\hline \mathrm{t}_s \\ \text{“Sunday”}(\mathrm{t}_s) \\ \mathrm{t}_s \subseteq \mathrm{t}_{ls} \\ \hline \end{array} \ \left\langle \begin{array}{c} \forall \\ \mathrm{t}_s \end{array} \right\rangle \ \begin{array}{|c|}\hline \mathrm{t}_{loc}\ \mathrm{e}\ \mathrm{t}_{adv} \\ \mathrm{on}(\mathrm{t}_{adv},\mathrm{t}_s) \\ \mathrm{t}_{loc} \prec \mathrm{n} \\ \mathrm{t}_{loc} \subseteq \mathrm{t}_{adv} \\ \mathrm{e} \subseteq \mathrm{t}_{loc} \\ \mathrm{e}:\ \text{“go-to-see”(m,a)} \\ \hline \end{array} \\ \hline \end{array}$$

representing the quantifications, and interpreting tense and outer adverb as constraints on this "state"; and (ii) treating the outer adverb as an additional restriction on the temporal quantification expressed by the PP. Neither of these mechanisms could account for the facts we observed without further assumptions, however, and even when the two are combined, extra assumptions are needed for which no compelling justification has yet been offered. We must leave the question of the interpretation of sentences like (112) in this unsatisfactory state, as an example of the many problems in this domain that are still waiting for a solution.

Frequency Adverbs

So far we have considered quantificational temporal adverbs which have the form of PPs in which the preposition governs a quantificational NP. The interpretation of frequency adverbs such as *always*, *often* etc. runs along much the same lines. But here we encounter additional complications. First, there is the problem how material within the scope of the adverb is to be divided between restrictor and nuclear scope. (cf. e.g. [Rooth, 1992] for the effects of information structure). This is a problem about which much has been and is being written, but it falls outside the scope of this survey. A second problem has to do with the interpretation of the contextual predicate C. In discussing quantificational PPs we focussed on the interpretation of C as inclusion (of the values of the bound variable) within a certain frame interval. With frequency adverbs this aspect of the interpretation of C is equally important. But in addition quantification by frequency adverbials is affected by another element of indeterminacy, which also can be contextually resolved or reduced, and often is. This second indeterminacy concerns the "granularity" of the quantification. For an illustration consider the following sentence:

(116) On Sunday Mary often called her aunt.

This sentence is ambiguous between an interpretation according to which there were many Sundays on which Mary called her aunt and a reading according to which there was a particular Sunday (e.g. the last one before the time on which the sentence was uttered) when Mary made many calls to her aunt. On the first reading the set of "cases" many of which are said to have been "cases when Mary called her aunt on Sunday" presumably consists of periods of the order of magnitude of a week. On the second reading the cases of which many are said to be "cases where Mary called" involve times of which a good many must fit within a single day. Part of what a speaker has to do when he has to assign meaning to sentences involing frequency adverbs is thus to form a conception of roughly what size periods are involved in the quantification it expresses. With nominal quantification the granularity question is normally resolved through the predication expressed by the nominal head of the quantifying phrase (cf. the noun *Sunday* in the quantifying NP of (103.d)), but with frequency adverbs granularity has to be determined by other means. For this reason the ambiguity we find in (116) is possible with the latter but not with the former. What general strategies are employed in arriving at granularity decisions when interpreting frequency adverbs is another question we can do no more than mention.

Negation

Sentence negation, as expressed in English by the word *not* (with or without do support), is among the operators of natural language which have a temporal and an aspectual dimension. As a rule, negation involves, implicitly or explicitly, some "frame" interval within which the negated condition is asserted to be unrealised. For instance, the statement

(117) Mary didn't call on Tuesday.

is understood as claiming that within the period denoted by *Tuesday* the condition of Mary calling did not obtain; in other words, that within this period there was no event of Mary calling. It is natural to associate with this observation the assumption that negation also has an aspectual effect, viz that irrespective of whether the material in its scope is stative or non-stative, the negated clause describes a state - a state to the effect that the given frame interval does not include an eventuality described by the clause to which the negation applies.

To capture these intuitions we consider the option of analysing negation as an aspect operator "NOT" which, like the operator "PROG", operates on properties of eventualities. The eventuality property is provided by the material in the scope of the negation – indeed, this perspective makes it natural to treat negation in a manner that is suggested by syntax for many of its actual occurrences - viz as a VP adjunct (nothing of the present proposal, though, really depends on this assumption.)

For the case of (117) the option gives rise to a representation of the following form:

$$(118)\quad \boxed{\begin{array}{c} \mathrm{t}\ \ \mathrm{t}'_{adv}\ \ \mathrm{m} \\ \mathrm{Mary(m)} \\ \mathrm{Tuesday}(\mathrm{t}'_{adv}) \\ \mathrm{t} < \mathrm{n}\ ;\ \mathrm{t} = \mathrm{dur(s)}\ ;\ \mathrm{t}'_{adv} = \mathrm{dur(s)} \\ \mathrm{s{:}NOT}(\hat{}\ \mathrm{e.e{:}call(m)}) \end{array}}$$

The conditions t=dur(s) and $\mathrm{t}'_{adv} = \mathrm{dur(s)}$ arise from the assumption that NOT has the properties of an adverbial quantifier and as such is subject to the same special constraints on the temporal location of the state it introduces as fequency adverbs like *always*, *often* or *never*.

(118) doesn't reveal much of the actual truth conditions associated with negation. This can be made more explicit via a meaning postulate for NOT, according to which the last condition of (118) can be written as in (119):

$$(119)\quad \boxed{\begin{array}{c} \mathrm{t} \quad \mathrm{t}'_{adv} \quad \mathrm{m} \\ \mathrm{Mary(m)} \\ \mathrm{Tuesday}(\mathrm{t}'_{adv}) \\ \mathrm{t} < \mathrm{n} \;;\; \mathrm{t} = \mathrm{dur(s)} \;;\; \mathrm{t}'_{adv} = \mathrm{dur(s)} \\ \mathrm{s}{:}\neg \boxed{\begin{array}{c} \mathrm{e} \\ \mathrm{e} \subseteq \mathrm{dur(s)} \\ \mathrm{e{:}call(m)} \end{array}} \end{array}}$$

In virtue of the condition dur(s)=t we can replace dur(s) by t. s has now become redundant. So we can eliminate all further occurrances of s, thus obtaining the reduced representation (120):

$$(120)\quad \boxed{\begin{array}{c} \mathrm{t} \quad \mathrm{t}'_{adv} \quad \mathrm{m} \\ \mathrm{Mary(m)} \\ \mathrm{Tuesday}(\mathrm{t}'_{adv}) \\ \mathrm{t} < \mathrm{n} \;;\; \mathrm{t} = \mathrm{t}'_{adv} \\ \neg \boxed{\begin{array}{c} \mathrm{e} \\ \mathrm{e} \subseteq \mathrm{t} \\ \mathrm{e{:}call(m)} \end{array}} \end{array}}$$

Although (120) is sufficient to capture the truth conditional content of (117), the alternative representations (118) and (119) are useful as well, in so far as they bring out the aspectual effect of negation and allow the rules which govern the temporal location of negation to be subsumed under the more general category of adverbial quantification. (120) should thus be considered as the result of a harmless simplification after a representation has first been constructed in the form given in (118), and then be transformed into (119) by application of the meaning postulate.

The present analysis brings out how negation can, through the ways in which it interacts with tense and temporal adverbs, produce an effect of temporal quantification. One consequence of this is that sentences containing negation expressed with the help of *not* often have the same truth conditions as sentences in which this negation is replaced by *never*. For instance (117) has the same truth conditions as

(121) On Tuesday Mary never called.

In fact, the two sentences may end up with the same semantic representations. Whether they do will depend on the exact treatment we adopt for the adverb *never*.

These proposals for treating negation are closely related to and in large part inspired by work of DeSwart (see [de Swart, 1999]).

Syntax and Semantics of the New Representation Language

Representating temporal information in the way which we argued to be necessary requires important extensions to the formalism which we had reached by the end of Section 3.1.[29] These extensions consist of

- new discourse referents for
  - points and periods of time,
  - events, and
  - states;
- a number of new predicates and functors in which entities of the sorts represented by the new discourse referents occur as arguments (as well as the atomic and non-atomic individuals exclusively considered hitherto). Among the predicates there are those which relate times and/or eventualities — the only ones we have had occasion to use here were $\prec$, $\subseteq$ and $\supset\!\subset$, but in general more are needed — as well as an open-ended number of predicates which relate eventualities to the entities of which we have been speaking throughout this chapter as (atomic and non-atomic) individuals. (These latter predicates, in which the eventuality argument is linked via a colon to the remainder of the predicational expression, are usually based on lexical verbs, although in the discussion of temporal quantification we also considered state predicates built from duplex conditions.)

From the point of view of predicate logic the new representation formalism is a system of many-sorted predicate logic. This is made explicit in both the syntax and the model theory for the new formalism which are given below. It is well-known that the transition from ordinary (1-sorted) predicate logic to many-sorted predicate logic is of little importance for metamathematics. Many-sorted formalisms can be embedded within their 1-sorted counterparts by adding predicates for the different sorts and adding postulates which express the sortal restrictions on the arguments of the original predicates and functors. It follows from this that systems of many-sorted first order logic are axiomatisable just as standard first order logic is; and

[29] Or alternatively, the extended formalisms of Sections 3.3 and 3.4. The extension described below is independent from those of 3.3 and 3.4.

the many-sorted variants of first order logic inherit other nice properties from standard first order logic as well. From this perspective the present extension thus seems much less dramatic than the introduction of plural discourse referents in the previous section.

However, our extensive experience with questions concerning the structure of time has taught us to be cautious. From a semantic perspective the present formalism is not just some arbitrary many-sorted generalisation of first order logic. It is a many-sorted logic the sorts of which are subject to certain conceptual constraints. For instance, time is conceived as a linear order, and some will go further and see its conception as carrying a commitment to its being unbounded, dense or even continuous (in the technical sense of being closed under limits of infinite ascending or descending sequences of bounded intervals). For someone who takes some or more of the sorts of the many-sorted system to be subject to such ontological constraints valid inference should mean "valid given that these constraints are satisfied". If this is the notion of "logical" validity we are after, the question whether a many-sorted system is axiomatisable can no longer be answered in a simple once-and-for-all manner. It now depends on the nature of the postulates which express the properties that are part of the ontological commitments. If these postulates are second order, then it may well be that validity ceases to be amenable to a finitary proof-theoretical characterisation.[30]

The problems we are facing when we pass from the DRT formalism of Section 3.1 to a many-sorted formalism where some of the sorts are assumed to come with a special structure are thus not unlike those which we encounter when we extend first order logic with non-standard quantifiers. What metamathematical properties our many-sorted system will have depends on what properties we assume for the different sorts it represents, just as the logical properties of extensions by non-standard quantifiers depend on the particular assumptions that are made about the semantics of those quantifiers.

We will not pursue these metamathematical questions here. (For the extensive knowledge that has been gathered about the effect of assumptions about the structure of time on the metamathematical properties of temporal

[30] The addition of plural discourse referents to the first order DRT formalism may be seen as a case in point. We could have introduced these as discourse referents of a new sort, whose values are sets (of cardinality ≥ 2; but this restriction has no importance in the present context). This would turn the representation system of the last section formally into a two-sorted system. That validity for this system cannot be axiomatised follows from the fact that the relationship between the values of the new discourse referents and those of the old ones – i.e. the relation that holds between sets and their members — is essentially second order. If we are content with less — e.g. by adopting one of the well-known first order set theories such as ZFU (Zermelo-Fränkel with Urelements), or GBU (Gödel-Bernays with Urelements) — as stating the relevant properties of and relations between the two sorts of individuals (viz. between Urelements and sets), then axiomatisability of validity is regained, albeit at the loss of the conceptually simplest way of conceiving of the realm of sets and its relation to the realm of things.

logics we refer to [Gabbay and Reynolds, 1994] and [Gabbay and Finger, 2000]. One of the general surprises within this domain has been no doubt that constraints on the structure of time which are irreducibly second order may nevertheless lead to notions of validity for temporal logics which are axiomatisable (or even decidable). (This is not always so, but it is true for a remarkably broad range of cases.) What metamathematical properties we get for the "first order part" of the formalism defined below (i.e. the part without plural discourse referents) on various assumptions about the structure of time is a question which to our knowledge has hardly been studied. We leave this as one of the many open problems of this section.

The DRS language we now proceed to define is to be regarded as a prototype. We have decided to include in it those symbols and expressions which make appearences in the DRSs that have been displayed in this section. A good deal more is needed for a representation language which is able to represent in a transparent and natural way all temporal information expressible by means of natural language devices.

Like the DRS language considered in Section 3.1 the vocabulary of the present one to define includes the following three categories of symbols.

(i) a set Ref of discourse referents,

(ii) a set Rel of predicates, and

(iii) a set Name of proper names.

In addition we allow for function symbols. In the language presented here this category plays only a marginal role. It contains only one element, viz. the functor DAY-OF which was used in the representation of *yesterday* in (92). However, in a full-blown DRS formalism for the representation of temporal information many more functors are needed. The same is true for the category of 1-place predicates of times. Many of these are calendar predicates — predicates such as *day*, *week*, *year* — which are true of a time iff it is a member of the various partitions of the time line into successive intervals which the calendar imposes on it. We have found use for one such predicate here, viz. DAY; but obviously that is one of a whole "network" of calendar concepts. (For more on the modeltheoretic semantics of calendar-predicates and other predicates which involve the metric of time, see [Kamp and Schiehlen, 2002].) Finally, we will make use of a 1-place predicate EXISTS in order to be able to represent contingency of existence.

As noted, the principal difference between the present DRS language and those introduced earlier is that the new one is many-sorted. This is reflected in the structure of the set Ref given in Definition 34.

DEFINITION 34. Ref is the union of the following four mutually disjoint sets of discourse referents.

Ind = $\{x_1, ..., x_n, ...\}$, a set of individual referents
Time = $\{t_1, ..., t_n, ...\}$, a set of referents for times
Event = $\{e_1, ..., e_n, ...\}$, a set of referents for events
State = $\{s_1, ..., s_n, ...\}$, a set of referents for states

We refer to the sets Ind, Time, Event and State as "sorts". We will later use the term "sort" also to refer to sets of entities in our models. No confusion should arise from this "overloading".

DEFINITION 35. The set Rel of relation symbols consists of

(i) n-place predicates of individuals;

(ii) $(n + 1)$-place predicates (with $n \geq 0$) where the first argument is an event and the remaining n arguments are of type individual (so-called $(n + 1)$-place event predicates);

(iii) $(n + 1)$-place predicates (with $n \geq 0$) where the first argument is of type state and the remaining n arguments are of type individual (so-called $(n + 1)$-place state predicates);

(iv) 2-place predicate symbols denoting temporal relations between times, events and states: $\prec$, $\subseteq$, $\supset\subset$;

(v) A 2-place predicate PROG, whose first argument is a state and whose second argument is a property of events;

(vi) A 1-place predicate of times: DAY;

(vii) A 1-place partial function of times: DAY-OF;

(viii) A 1-place predicate of individuals, events and states: EXISTS;

(ix) A 1-place functor from eventualities to times: DUR;

As before, DRS-conditions and DRSs are defined by simultaneous recursion. In Definition 36 we only specify the new clauses of the definition; they should be seen as supplementary to those of Definition 2. ((ix) replaces the earlier clause 2.ii for conditions of the form "$x_i = x_j$".)

DEFINITION 36. DRS conditions:

(i) if $\tau, \sigma \in$ Event $\cup$ State $\cup$ Time, R one of the predicates $\prec$, $\subseteq$ and $\supset\subset$, then τ R σ is a condition;

(ii) if e $\in$ Event, x_1,..,x_n $\in$ Ind and R $\in$ Rel an $(n + 1)$-place event predicate, then e:R(x_1,..,x_n) is a condition;

(iii) if s $\in$ State, x_1,..,x_n $\in$ Ind and R $\in$ Rel an $(n+1)$-place state predicate, then s:R(x_1,..,x_n) is a condition;

(iv) if s $\in$ State, e $\in$ Event, K a DRS and e $\in U_K$, then s:PROG($^\wedge$ e.K) is a condition;

(v) if s $\in$ State, t $\in$ Time, K_1 and K_2 are DRSs and t $\in U_{K_1}$, then s:K_1 $\left\langle \overset{\forall}{\underset{t}{}} \right\rangle$ K_2 is a condition;

(vi) if t $\in$ Time, then DAY(t) is a condition;

(vii) if t_1, t_2 $\in$ Time, then t_1 = DAY-OF(t_2) is a condition;

(viii) if $\tau \in$ Event $\cup$ State $\cup$ Ind, then EXISTS(τ) is a condition.

(ix) if τ and σ are discourse referents of the same sort, then $\tau = \sigma$ is a condition.

The model theory for the DRS language defined above raises a number of fundamental questions. Some of these concern status and structure of the ontological categories of times, events and states, the relations between them and the relations between them and the category of (atomic and non-atomic) individuals which have been the sole denizens of the models considered up until this point. Secondly, there is the problem of contingent existence, which was mentioned briefly in Section 3.2 in connection with intensional models. In fact, in the present context this problem arises twice over, once in connection with possible worlds — what exists need not have existed necessarily – and once in connection with time — what exists at one time need not exist at every other time. Finally, models which involve both worlds and times raise the question how worlds and times are connected. An important part of our conception of possibility and necessity has to do with future contingency: our actual world can develop into one future or another, so what is one world at one time may turn onto one of a number of different possible worlds at a later time.

We begin with the problems which concern the ontological status of times and eventualities, their structural properties and the relations between them and individuals. The first question that an ontologist is likely to ask about these or any of the categories is in what sense, if any, entities belonging to them "exist", or are "real". Here the question is a fairly ramified one, since (apart from the category of individuals of which we will assume for simplicity's sake that the question has already been answered) we are dealing with three categories at once — times, events and states. So a whole range of possible answers is possible in principle. One possible position is that only times constitute a primitive domain of "irreducible existents" and that events and states constitute "virtual" or "derived" entities which should be seen as constructs out of times (in combination, presumably, with entities from other sorts, such as individuals, properties or relations). But a diametrically opposed position, according to which events form an irreducible category and times are constructions out of events, has been put forward also (with or without the supplementary assumption that states are constructs defined from this basis as well). Yet another position is the one according to which events are to be analysed as transitions between states, and thus that the category of events is reducible to the category of states. (For discussion of some of these alternatives see [Benthem, 1983; Kamp and Reyle, 1993] and references there.)

This list is surely not exhaustive. But it suffices to show that the model theory for the DRS language we have specified in Definitions 34-36 might be grounded in a number of different ways, and that the philosophical logician is likely to prefer one version or another depending on his metaphyiscal persuasions. In the model theory we develop here we remain neutral on these issues of ontological priority and reducibility. Note, however, that we are committed to models in which all of the four mentioned sorts — individuals, times, events and states — are represented. For the vocabulary of our DRS language includes discourse referents of each of these sorts, and we want to stick to the general form of our semantic definitions, all of which are based on the notion of an assignment which maps discourse referents onto suitable entities in the model. In the context of this section (as in that of the last section) this entails that a discourse referent belonging to any one of these sorts should be always assigned entities of the model which are of its own sort. Under these constraints neutrality on matters of ontological reducibility can only mean this:

> The universe of a model $\mathcal{M}$ is composed of the four categories $\mathrm{Time}_{\mathcal{M}}$, $\mathrm{Event}_{\mathcal{M}}$, $\mathrm{State}_{\mathcal{M}}$, $\mathrm{Individual}_{\mathcal{M}}$. Whether any one of these categories can be reduced to any combination of the others is left open. Models which involve such reductions are not excluded. But they will be only some among the totality of all models admitted by the general definition we will give.

As far as the time structure of our models is concerned we want to be very specific. We are persuaded that people's intuitions about the structure of the time of the external world are, when pushed hard enough, to the effect that time is like the real numbers; so we will assume that the time structure $\mathcal{T}_{\mathcal{M}} = \langle T_{\mathcal{M}}, \prec_{\mathcal{M}} \rangle$ of $\mathcal{M}$ is isomorphic to the reals. Our assumptions about the structure of events and states are less specific. The times of the model $\mathcal{M}$ that are targets for assignments to the discourse referents in $\mathrm{Time}_{\mathcal{M}}$ are not the "points of time" which make up the set $\mathcal{T}_{\mathcal{M}}$, but the "intervals" which can be formed out of these. The notion of interval must be handled with some care, however, since the distinction between open and closed intervals of $\mathcal{T}_{\mathcal{M}}$ is meaningless from the perspective of natural language interpretation. We can eliminate the open-closed distinction either by forming equivalence classes of convex subsets of T — for two such sets X and Y we put $X \equiv Y$ iff $Cl(X) = Cl(Y)$, where $Cl(X)$, the "closure of X", is the set consisting of all limits of converging sequences of points in X — or, alternatively, by taking unique representatives of the equivalence classes of $\equiv$, for instance the intervals $(t_1, t_2]$ with $t_1 \prec_{\mathcal{M}} t_2$, together with $(-\infty, t_2]$, (t_1, ∞) and $(-\infty, \infty)$. These two options are not fully equivalent in sofar as the first includes the points $t \in T_{\mathcal{M}}$ themselves – in the form of singleton equivalence classes $\{[t]_{\equiv}\}$ – whereas the second leaves them out. (There is no such half-open, half-closed interval as $(t, t]$.) In connection with the DRS language of Definitions 34-36 this difference appears to be of no importance but for definiteness' sake we arbitrarily choose the second option. We refer to this set of intervals of $\mathcal{T}$ as Time($\mathcal{T}_{\mathcal{M}}$). (In connection with certain richer representation languages the question whether "intervals" consisting of single points should be included gains importance and must be considered carefully.)

We assume that each model $\mathcal{M}$ has a set $E_{\mathcal{M}}$ of events and a set $S_{\mathcal{M}}$ of states, that these sets are disjoint and that together they form the set of eventualities $EV_{\mathcal{M}}$ of $\mathcal{M}$. $EV_{\mathcal{M}}$ is part of an eventuality structure $\langle EV_{\mathcal{M}}, \prec_{\mathcal{M}}, \bigcirc_{\mathcal{M}} \rangle$, which is assumed to satisfy the following postulates.

DEFINITION 37. An *eventuality structure* $\mathcal{EV}_{\mathcal{M}}$ is a triple $\langle EV_{\mathcal{M}}, \prec_{\mathcal{M}}, \bigcirc_{\mathcal{M}} \rangle$ with $EV_{\mathcal{M}} = E_{\mathcal{M}} \cup S_{\mathcal{M}}$ where $E_{\mathcal{M}}$ is a set of events and $S_{\mathcal{M}}$ a set of sates. $\mathcal{EV}_{\mathcal{M}}$ satisfies for all eventualities $ev, ev_1, ..., ev_4 \in EV_{\mathcal{M}}$:

(1) $(ev_1 \prec_{\mathcal{M}} ev_2) \rightarrow \neg(ev_2 \prec_{\mathcal{M}} ev_1)$
(2) $(ev_1 \prec_{\mathcal{M}} ev_2 \wedge ev_2 \prec_{\mathcal{M}} ev_3) \rightarrow (ev_1 \prec_{\mathcal{M}} ev_3)$
(3) $ev \bigcirc_{\mathcal{M}} ev$
(4) $(ev_1 \bigcirc_{\mathcal{M}} ev_2) \rightarrow (ev_2 \bigcirc_{\mathcal{M}} ev_1)$
(5) $(ev_1 \prec_{\mathcal{M}} ev_2) \rightarrow \neg(ev_2 \bigcirc_{\mathcal{M}} ev_1)$
(6) $(ev_1 \prec_{\mathcal{M}} ev_2 \wedge ev_2 \bigcirc_{\mathcal{M}} ev_3 \wedge ev_3 \prec_{\mathcal{M}} ev_4) \rightarrow (ev_1 \prec_{\mathcal{M}} ev_4)$
(7) $ev_1 \prec_{\mathcal{M}} ev_2 \vee ev_1 \bigcirc_{\mathcal{M}} ev_2 \vee ev_2 \prec_{\mathcal{M}} ev_1$

$\mathcal{EV}_{\mathcal{M}}$ and $\mathcal{T}_{\mathcal{M}}$ are correlated via a function LOC$_{\mathcal{M}}$ which maps the eventualities in $EV_{\mathcal{M}}$ onto intervals of $\mathcal{T}_{\mathcal{M}}$, thereby locating these eventualities on the time axis defined by $\mathcal{T}_{\mathcal{M}}$. Thus LOC$_{\mathcal{M}}$ is assumed to assign each $ev \in EV_{\mathcal{M}}$ an interval in Time($\mathcal{T}_{\mathcal{M}}$). We assume that LOC$_{\mathcal{M}}$ preserves the temporal relations of $\mathcal{EV}_{\mathcal{M}}$, that is: if $ev_1, ev_2 \in EV_{\mathcal{M}}$, then

- if $ev_1 \bigcirc_{\mathcal{M}} ev_2$, then $\mathrm{LOC}_{\mathcal{M}}(ev_1) \cap \mathrm{LOC}_{\mathcal{M}}(ev_2) \in Time(\mathcal{T}_{\mathcal{M}})$,
- if $ev_1 \prec_{\mathcal{M}} ev_2$, then $\mathrm{LOC}_{\mathcal{M}}(ev_1) \prec_{int} \mathrm{LOC}_{\mathcal{M}}(ev_2)$
(where $\prec_{int}$ is the relation which holds between two intervals $(\mathrm{t}_1, \mathrm{t}_2]$ and $(\mathrm{t}'_1, \mathrm{t}'_2]$ in Time$_{\mathcal{M}}$ iff $\mathrm{t}_2 \prec_{\mathcal{M}} \mathrm{t}'_1$.[31]

[31]It is well known that structures $\mathcal{EV}$ satisfying the conditions of Definition 37 give rise

So far we have identified as components of our models:

1. A time structure $\mathcal{T}_{\mathcal{M}}$
2. An eventuality structure $\mathcal{EV}_{\mathcal{M}}$
3. An embedding $\mathrm{LOC}_{\mathcal{M}}$ of the latter in the former

What we need in addition are:

4. A universe $\mathrm{U}_{\mathcal{M}}$ of individuals
5. Interpretations for the predicates of the DRS language (specified in Definition 35)

Among the predicates there are three structural "predicates", viz $\prec$, $\subseteq$ and $\supset\!\subset$, whose interpretation is determined by the information provided in 1-3 above. For instance, the interpretation $\Im_{\mathcal{M}}(\prec)$ is defined as follows:

---

to an instant structure $\mathcal{I}(\mathcal{EV}) = \langle I(\mathcal{EV}), \prec (\mathcal{EV})\rangle$, where

– $I(\mathcal{EV})$ consists of all maximal sets of pairwise overlapping members of $\mathcal{EV}$ (i.e. $i \in \mathcal{I}(\mathcal{EV})$ iff (i) $i \subseteq EV$, (ii) whenever $ev_1, ev_2 \in i$, then $ev_1 \bigcirc ev_2$, and (iii) if $i \subseteq H \subseteq EV$ and H has the property that $ev_1 \bigcirc ev_2$ whenever $ev_1, ev_2 \in H$, then $H \subseteq i$).

And

– for $i_1, i_2 \in I(\mathcal{EV})$, $i_1 \prec_i i_2$ iff there are $ev_1 \in i_1$ and $ev_2 \in i_2$ such that $ev_1 \prec ev_2$.

On the basis of these definitions it is easy to show that $\mathcal{I}(\mathcal{EV})$ is a linear order, that for each $ev \in EV$ the set if $i \in I(\mathcal{EV})$ such that $ev \in i$ forms a convex subset of $I(\mathcal{EV})$, and that the relation "$ev \in i$" is naturally interpreted as saying the i is a period of time at which ev is going on.

We might expect that the function $\mathrm{LOC}_{\mathcal{M}}$ induces an order preserving embedding LOC' of $\mathcal{I}(\mathcal{EV})$ into the interval structure $\mathcal{INT}(\mathcal{T}_{\mathcal{M}})$ of $\mathcal{T}_{\mathcal{M}}$ via the condition

(i) $\mathrm{LOC}'(i)$ is that non-empty interval $(t_1, t_2]$ such that $(t_1, t_2] = Cl(\bigcap\{\mathrm{LOC}_{\mathcal{M}}(ev) | ev \in i\})$,
where for arbitrary $X \subseteq T$ $Cl(X)$ denotes the convex hull of X in $\mathcal{T}$.

However, in general this need not be so. On the other hand, if LOC' is such an embedding then LOC can conversely be defined in terms of it via

(ii) $\mathrm{LOC}_{\mathcal{M}}(ev) = (t_1, t_2]$, where $(t_1, t_2] = CL(\cup\{\mathrm{LOC}'(i) : ev \in i\})$

More generally, when LOC' is any order preserving map from $\mathcal{I}(\mathcal{EV}_{\mathcal{M}})$ into $\mathcal{T}_{\mathcal{M}}$ and a function LOC on $EV_{\mathcal{M}}$ is defined from LOC' via (*), then LOC is order preserving.

We conclude that in some models $\mathcal{M}$ $\mathrm{LOC}_{\mathcal{M}}$ will be derivable from an underlying map LOC', but not in all.

DEFINITION 38.
(a) Let $\alpha \in \mathrm{Time}(\mathcal{T}_{\mathcal{M}}) \cup EV_{\mathcal{M}}$, then

$$\mathrm{Time}(\alpha) = \begin{cases} \alpha, & \text{if } \alpha \in \mathrm{Time}(\mathcal{T}_{\mathcal{M}}) \\ \mathrm{LOC}_{\mathcal{M}}(\alpha), & \text{if } \alpha \in EV_{\mathcal{M}} \end{cases}$$

(b) $\Im_{\mathcal{M}}(\prec) = \{ \langle \mathrm{Time}(\alpha), \mathrm{Time}(\beta) \rangle |$
$\alpha, \beta \in \mathrm{Time}(\mathcal{T}_{\mathcal{M}}) \cup EV_{\mathcal{M}} \wedge \mathrm{Time}(\alpha) \prec_{int} \mathrm{Time}(\beta) \}$

(c) The definitions of $\Im_{\mathcal{M}}(\subseteq)$ and $\Im_{\mathcal{M}}(\supset\subset)$ are left to the reader.

The extension of the predicate DAY in $\mathcal{M}$ should partition $T_{\mathcal{M}}$ into a set of intervals which is order-isomorphic to some subset of the integers (some subset of the integers rather than all of the integers, since we want to allow for the possibility that there is a first and/or a last day).

Once the interpretation $\Im_{\mathcal{M}}(\mathrm{DAY})$ is given, this also fixes the interpretation of the partial functor DAY-OF: If t is an interval belonging to $Int(\mathcal{T}_{\mathcal{M}})$ and there is a member d of $\Im_{\mathcal{M}}(\mathrm{DAY})$ such that $t \subseteq d$, then DAY-OF(t) is d. Otherwise DAY-OF is undefined.

We will not impose any constraints on the other predicates of Definition 35 (except for the predicates PROG and EXISTS, to which we will come below). In order to obtain a "realistic" class of models many further constraints would be desirable. However, formulating such constraints is a notoriously difficult problem.

The second problem about which something needs to be said is that of contingent existence. In relation to the models that are needed here this problem arises "twice over", we noted, once in connection with time and once in connection with modality. From a general logical point of view the problem is the same in either case; it constitutes one part of what in the classical analytical literature on modality is known as the problem of "quantifying in" (See among others: [Quine, 1961; Quine, 1956; Kaplan, 1969]). In DRT-terminology the problem can be described as follows. (We give the description for the case of worlds, but the version for times is analogous.) Suppose that in the process of evaluating a DRS K or a DRS condition γ in a given world w we assign to a given discourse referent x an entity d which exists in w, and suppose that the structure of K or γ requires that we evaluate parts of it which contain free occurrences of x at some other world in which d does not exist. In this case the evaluation will abort, and it is quite possible that it will abort for what is intuitively the wrong reason. A truth definition which does not handle this problem with the care it requires is likely to create a lot of truth value gaps in places where there shouldn't be any.

Since the contingent existence problem arises as much in relation to time as in relation to possible worlds, the model theory for our present DRS language would have to deal with it even if it were kept purely extensional. But since what we want is an intensional model theory, we have to address both the temporal dimension of it and the possible world dimension. As a matter of fact we will not really deal with either dimension of the problem, but follow the avoidance strategy we adopted in Section 3.2: we blithely assume that everything that exists exists both necessarily and eternally. This formally avoids the quantifying-in problem we have described, but at the price of a notion of model that is blatantly unrealistic. However, in applications to the semantic analysis of natural language the conceptual disadvantages of this crude simplification can be minimised through the judicious use of existence predicates — predicates the extension of which at a given time in a given world consists of what exists at that time in that world. The extensions of such predicates will normally vary as a function of both worlds and times. By inserting existence predicates into the semantic representations of sentences or discourses the most nefarious manifestations of the quantifying-in problem can usually be avoided. In formulating the satisfaction conditions for the DRS language under consideration we will encounter one problem for whose solution we will need an existence predicate. For this reason we add such a predicate to our language. We denote it as "EXISTS". The contingency of existence which EXISTS allows us to represent is limited: it accounts for variation between worlds, but not between times within the same world. (In order to account for variation between times as well within the present formalism an existence predicate would have to be a 2-place predicate with an additional argument for times. Since variation between times is not needed for the application alluded to, we have decided to make do with the simpler version of a 1-place predicate.)

As we have already made the decision to adopt a notion of model which sweeps the problem of contingent existence under the rug, further discussion of this problem may seem an unwanted luxury. However, we want to point at some of the more specific problems that will have to be dealt with by a model theory in which the contingency problem is taken seriously. In particular we want to draw attention to the fact that behind the superficial similarity we have noted between the temporal and the modal dimension of the problem hide what seem to be important differences. One is that in the case of time an important role is played by temporal order: once something has existed, it continues to be something that can be referred to (for instance in order to assert of it that it exists no longer); but it is dubious whether something can be an object of reference at a time before it comes into existence. This contrast seems to be particularly pronounced for eventualities: for an event or state there is the time at which it happens or holds. But it is entirely natural to refer to it at later times as something that did occur or hold at the earlier time. (In fact, in almost all cases where we have made use of eventuality discourse referents in the DRSs above the discourse referents play just this role: they serve to represent events or states of which the DRS claims that they occurred at some time distinct from the utterance time.)

A special case for the question of contingent existence are the times themselves. Were we to assume that times "exist only at themselves", and could not be referred to at any other time, then meaningful talk about time and times would be impossible. If we are to acknowledge time as an ontological category at all, then only as one whose elements are possible subjects of discussion at all times. In other words, times must — paradoxical as that may sound — be eternal if they are to be anything at all.

This doesn't settle the modal dimension of the existence of time. We may still ask: is the time structure of one possible world the same as that of another, or could they be different? This is a question which is closely connected with the problem of ontological priority we mentioned earlier. Someone who sees time as an invariant receptacle within which the contingencies of the actual world unfold in the way they happen (whether this receptacle is to be seen as a metaphyiscal given in the sense of Newton or as a cognitively necessary condition on experience in the sense of Kant) will be inclined to assume that time is the same in all possible worlds. Someone who sees time as an epiphenomenon generated by the actual course of events and whose structure is a reflection of the underlying event structure, would expect the structure of time to vary from world to world — like the underlying courses of events on which it depends.

In the light of these possibilities, our assumption that time is necessarily isomorphic to the reals reveals a definite *parti pris*. It is an assumption which reflects our conviction that what matters in a model-theoretic treatment of meaning in natural language is our conception of time, which informs the ways in which we think and speak. However, by itself the claim that time is necessarily isomorphic to the reals doesn't determine whether all worlds of a given model have the same time. Two worlds w_1 and w_2 could each have a time structure isomorphic to the reals and yet the set of times of w_1 might be disjoint from the set of times of w_2. If that were so, there would be no natural way of comparing the times of w_1 with those of w_2 — there would be no straightforward way of "synchronising" the two worlds. In particular there would be no way of determining which time in w_2 corresponds to a time t_u at which a certain utterance is made in w_1. This is a situation that, in the light of what we need our model theory for, should be avoided. The simplest (and most radical) way to avoid it is to assume that all worlds of a given model $\mathcal{M}$ have one and the same time structure $\mathcal{T}_{\mathcal{M}}$; and this is what we do. (In fact, we already made this decision, since it is entailed by the more general one according to which all four ontological categories are constant between the different worlds of a given model.)

The last of the problems we mentioned above has to do with future contingency. In the philosophical literature this problem has often been discussed under the heading of "historical necessity" — a proposition about the future is historically necessary at a given point in time t iff it is necessarily true in virtue of what has been the case up to t and is the case at t itself.

A natural way of modelling the intuition that some of the things that will happen later will happen as a matter of historical necessity while others will happen contingently, is as follows: a given world w, as it has developed up to the time t, can go on after t in any one of a number of different ways; these different ways form a "bundle" of future continuations of w after t which between them cover all that is possible in the light of what is and has been the case in w at t. It is common to formalise this by means of a 3-place relation between two worlds w_1 and w_2 and a time t, a relation which holds between w_1, w_2 and t iff w_1 and w_2 are alternative possible continuations of what was still a single world at t.

This relation between worlds and times has proved indispensible to the semantic and logical analysis of a significant range of natural language expressions and constructions. (And the same is true for a number of aspects of the interactive structure of worlds and times). Should one want to use the model theory developed here in the analysis of any of these, then it will have to be refined by endowing its models with additional structure of the kind discussed (see [Thomason, 2002]). In connection with the DRS language we have defined here, however, the additional structure wouldn't do any work. So it isn't mentioned in the definition of models below.

DEFINITION 39. An *(intensional) model* $\mathcal{M}$ for the DRS language specified in Definitions 34-36 is a tuple $\langle \mathrm{W}, \mathrm{U}, \mathcal{EV}, \mathcal{T}, \mathrm{LOC}, \Im \rangle$, where W is a non-empty set of worlds, U a non-empty set of individuals, where $\mathcal{EV}$, $\mathcal{T}$ and LOC are described as above, and $\Im$ is a function which assigns to each non-logical constant of our DRS language an appropriate extension at each world $w \in \mathrm{W}$ and is subject to the constraints expressed in Def. 38 and those mentioned in the three paragraphs following it.

We have already assumed that the universe of individuals U, the eventuality structure $\mathcal{EV}$ and the time structure $\mathcal{T}$ are the same for all worlds of $\mathcal{M}$. What about LOC? Here we do want to allow for variation. The intuition is that the same eventuality could have happened earlier in one world than it did in another, or that it could have taken more or less time in the first world than in the second. We achieve this by allowing LOC(ev) to be different intervals in different worlds. So we assume that LOC is not simply a function from $\mathcal{EV}$ to Time($\mathcal{T}_{\mathcal{M}}$), but that it maps the worlds of W to such functions.

With regard to the interpretation function $\Im$ the question of variability arises as well. We assume that the interpretation of the following non-logical constants of our DRS language are rigid (i.e. that they do not vary).

(i) the proper names of our language, i.e. the members of Name,

(ii) the relations $\subseteq$, $\supset\!\subset$ and $\prec$,

(iii) the predicate DAY and the functor DAY-OF.

In other words, for each such expression α from this list we stipulate that if w, w' are any two worlds from W, then $\Im(\alpha)(w) = \Im(\alpha)(w')$.

For Name the assumption of rigidity was already made in Section 3.2, for DAY it is a stipulation for which we take the motivation to be clear, and for $\prec$, $\subseteq$, $\supset\!\subset$ and (given the rigidity of DAY) DAY-OF it follows from the definitions given above.

For all other non-logical constants we assume that they are not rigid. For the predicates some or all of whose arguments are of the sort individual we take this to need no justification. Likewise for the predicate PROG. For the functor dur non-rigidity is a consequence of the non-rigidity of LOC together with the self-evident principle that dur should be interpreted as LOC — that is, for every $w \in$ W $\Im(\text{dur})(w)$ is the function which maps each $ev \in EV$ onto LOC$(w)(ev)$. The non-rigidity of EXISTS is of course the very point of the presence of this predicate in our language.

It should be emphasised once more that non-rigidity of the sort allowed for in our models does not give us as much variation as one might want. In particular, it fails to account for the temporal variation of predicates which in natural language are expressed by means of nouns, adjectives and prepositions. Such natural language predicates typically vary their extensions over time (being *blond, under 65 kg*, or *a student* are properties which a person may have at one time without having them at all times.) The n-place predicates between individuals, which are intended as the formal representatives of such non-verbal natural language predicates do not capture this dimension of variation. One way to deal with this problem is to represent non-verbal n-place predicates of natural language by means of $(n+1)$-place state predicates where necessary and keeping the n-place predicates of Definition 36 only for those natural language predicates which are "eternal" in the sense that when an individual (or tuple of individuals) satisfies it at one time, it satisfies them at all times.

Most of what needs to be said towards the definition of truth and other semantic relations between expressions of our DRS language and models has been said already. The new DRS conditions are, with only a couple of exceptions, simple atomic conditions for which the satisfaction conditions they contain are determined directly by the interpretations assigned to the non-logical constants they contain. One example should be enough to establish the general pattern. We choose conditions of the form "e:R(x_1,...,x_n)". In Section 3.2 we showed various forms in which notions like satisfaction and truth can be defined. Here we focus on the first of these, according to which an assignment verifies a DRS condition in a model at a world: $g \models_{\mathcal{M},w} \gamma$ (see Definition 19 of Section 3.2). On the basis of these satisfaction clauses we can then define all other semantic notions introduced in 3.2 along the lines given there.

In this format the satisfaction clause for a condition of the form e:R(x_1,..., x_n) takes the following form:

Let $\mathcal{M}$ be a model in the sense of Definition 39, $w \in$ W and g an assignment which maps e onto an element of EV and x_1, ..., x_n onto elements of U. Then

(122) $g \models_{\mathcal{M},w} \mathrm{e{:}R}(\mathrm{x}_1,...,\mathrm{x}_n)$ iff $\langle g(\mathrm{e}), g(\mathrm{x}_1), ..., g(\mathrm{x}_n)\rangle \in \Im(\mathrm{R})(w)$

Of the remaining DRS conditions listed in Definition 36 there are three which need special attention. The first and easiest of these are DRS conditions of the form "t = DAY-OF(t′)". We defined $\Im$(DAY-OF) as a partial function from intervals of Time($\mathcal{T}$) to intervals of Time($\mathcal{T}$) which is defined only if the argument is included in a member of $\Im$(DAY). Partiality doesn't lead to truth value gaps in this case, because of the fact that terms of the form "DAY-OF(t′)" only occur in the context of conditions of the form "t = DAY-OF(t′)". The following obvious satisfaction condition makes this clear:

Assume that $\mathcal{M}, w$ are as above and that g maps t and t′ to members of Time($\mathcal{T}$). Then

(123) $g \models_{\mathcal{M},w} \mathrm{t} = \mathrm{DAY\text{-}OF}(\mathrm{t}')$ iff $\mathrm{t} \in \Im(\mathrm{DAY})$ and $\mathrm{t}' \subseteq \mathrm{t}$.

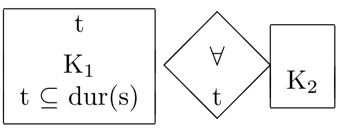
The second clause that deserves attention is that for conditions of the form "s:PROG(^e.K)". Actually the satisfaction conditions follow the pattern of (122):

(124) $g \models_{\mathcal{M},w} \mathrm{s{:}PROG}(^\wedge\mathrm{e.K})$ iff $\langle g(\mathrm{s}), [\![^\wedge\mathrm{e.K}]\!]_{\mathcal{M}}\rangle \in \Im(\mathrm{PROG})(w)$

We mention conditions of this form nevertheless because they contain – as the only ones among all the atomic conditions of our DRS language — terms which are not simply discourse referents. These terms are the property terms that occur as second arguments of PROG. The presence of these terms provides no real obstacles to our truth definition. But the fact that they don't is something which deserves explicit notice. For it is here that, for the first time, our choice of an intensional model theory for the present DRS language proves to be essential. In view of the developments in Section 3.2 the definition of satisfaction and truth of which (122) and (123) are constitutive clauses yields among other things a denotation for terms of the form ^e.K. For this reason we can assume the property $[\![^\wedge\mathrm{e.K}]\!]_{\mathcal{M}}$ to be defined at the point where it is needed in the definition of the satisfaction condition of "s:PROG(^e.K)".

The last and most problematic type of DRS condition is that which uses duplex conditions to characterise states as quantificational states. We repeat the general form of such conditions in (125).

(125) s:



(N.B. the box on the left should be seen as follows: it is a DRS K such that $t \in U_K$ and $t \subseteq dur(s) \in Con_K$.)

The problem with these conditions is that so far we have done no more than hint at what truth conditions they represent. We have described the quantificational state s as one that is to the effect that the quantification holds over the period of its duration. But what exactly does this mean and how could it be made precise? We propose the following: in order that s be a state to the effect that the given quantification holds over the period that it defines, the proposition that s exists must be the same as the proposition that the quantification holds over the given period. This leads for conditions of the form (125) to the satisfaction condition in (126). (It is at this point, and at this point only, that we have to make use of our existence predicate EXISTS in the satisfaction and truth definition of our DRS language.)

$$(126)\ g \models_{\mathcal{M},w} \mathrm{s}\text{:}\ \boxed{\begin{array}{c} \mathrm{t} \\ \mathrm{K}_1 \\ \mathrm{t} \subseteq \mathrm{dur(s)} \end{array}} \left\langle \begin{array}{c} \forall \\ \mathrm{t} \end{array} \right\rangle \boxed{\mathrm{K}_2} \quad \text{iff}$$

there is an interval $t_{\mathrm{fr}} \in \mathrm{Time}(\mathcal{T}_{\mathcal{M}})$ such that $t_{\mathrm{fr}} = [\![\mathrm{dur(s)}]\!]_{\mathcal{M},g}$ and

$$(*)\ [\![{}^\wedge\mathrm{EXISTS(s)}]\!]_{\mathcal{M},g} = [\![{}^\wedge \boxed{\boxed{\begin{array}{c} \mathrm{t} \\ \mathrm{K}_1 \\ \mathrm{t} \subseteq \mathrm{dur(s)} \end{array}} \left\langle \begin{array}{c} \forall \\ \mathrm{t} \end{array} \right\rangle \boxed{\mathrm{K}_2}}\ ]\!]_{\mathcal{M},g'},$$

where $g' = g \cup \{\langle \mathrm{t}', t_{\mathrm{fr}} \rangle\}$.

N.B. In general there is no reason to assume that the condition (*) in (126) determines s uniquely. But the idea that s is exhaustively characterised by this condition is not all that far-fetched; and it would be possible to adopt the condition that this is so as a general constraint on models.

This completes the satisfaction definition in essence. The complete definition is obtained by combining (123), (124), (126) with (a) clauses for the other atomic conditions of Definition 36 for which (122) serves as example, and (b) the clauses of the Satisfaction Definition 19 of Section 3.2. As we already observed, all the other semantic relations mentioned in 3.2 can be defined for the extended languages too.

To conclude, a remark relating to the DRS conditions (iv) and (v) of Definition 36. We begin with the PROG-condition defined in (iv). We argued that the existence of an event which statisfies a DRS K is not a neccessary condition for the existence of a state s such that s:PROG(^e.K). But intuitively the condition is sufficient and if it is to be that also formally, then there should be enough states around to make it so. In order to make sure of this we must impose on our models the requirement that they verify the following existence postulate.[32]

$$(127)\quad \boxed{\begin{array}{c} \mathrm{e\ t} \\ \mathrm{K} \\ \mathrm{t} \subseteq \mathrm{e} \end{array}} \quad \Rightarrow \quad \boxed{\begin{array}{c} \mathrm{s} \\ \mathrm{t} \subseteq \mathrm{s} \\ \mathrm{s{:}PROG}({}^\wedge\mathrm{e.K}) \end{array}}$$

(Here we have followed the same convention as in (125): The box on the left hand side of $\Rightarrow$ is a DRS K such that $\mathrm{e} \in \mathrm{U_K}$ and "$\mathrm{t} \subseteq \mathrm{e}$" $\in \mathrm{Con_K}$.)

The quantificational state conditions specified in Definition 36.v also cry out for a supporting existence postulate. In this postulate we make use of the same principle which we also used in defining the satisfaction condition of quantificational state conditions: if a temporal quantification condition holds over a period of time $\mathrm{t_{fr}}$ then there exists a state the duration of which is $\mathrm{t_{fr}}$ and which exists in any world w iff the quantificational state condition holds over $\mathrm{t_{fr}}$ in w.

$$(128)\quad \boxed{\begin{array}{c} \mathrm{t}_{fr} \\ \boxed{\begin{array}{c} \mathrm{t} \\ \mathrm{K_1} \\ \mathrm{t} \subseteq \mathrm{t}_{fr} \end{array}} \left\langle \begin{array}{c} \forall \\ \mathrm{t} \end{array} \right\rangle \boxed{\begin{array}{c} \mathrm{s} \\ \mathrm{dur(s)} = \mathrm{t}_{fr} \\ {}^\wedge\mathrm{EXISTS(s)} = {}^\wedge \boxed{\boxed{\begin{array}{c} \mathrm{t} \\ \mathrm{K_1} \\ \mathrm{t} \subseteq \mathrm{t}_{fr} \end{array}} \left\langle \begin{array}{c} \forall \\ \mathrm{t} \end{array} \right\rangle \boxed{\mathrm{K_2}}} \end{array}} \end{array}}$$

(127) and (128) can be regarded as meaning postulates. Meaning postulates play the same role in the model theory of DRS languages as they do in Montague Grammar; they act as constraints on models which narrow the class of models down by eliminating models which violate the semantic adequacy conditions they express. In this regard (127) and (128) do not differ, of course, from the axioms on $\mathcal{EV}$ given in Definition 37, the postulate that $\mathcal{T}$ be isomorphic to the reals, or the conditions of Definition 31 in Section 3.4 which articulate the mereological structure of U. However, the bulk of meaning postulates that will be needed to arrive at a satisfactory

[32]It may be felt that this is not quite right in so far as the progressive state does not hold up to the very end of e. To formulate the meaning postulate in a way that takes account of this we would need a richer vocabulary for expressing temporal relations than the given DRS language provides. Another possible objection against (127) is that it is wrong for the progressives of so-called achievement verbs such as *die*. *He was dying* expresses a state which is usually seen as preceding the event of death itself, rather than as being included in it. To deal with this, (127) should either be replaced by a weaker disjunction which distinguishes between achievements and accomplishments, or else one would have to assume that the interpretation of sentences like *He was dying* involves as an intermediate step extending the predicate *die* to one which is true of events that the process that leads up to the actual death is an integral part of.

model theory for a DRS language suitable for the representation of natural language have to do with the meanings of individual lexical items such as nouns and verbs. We will consider some examples of such postulates in a forthcoming article.

The DRS language for which we defined syntax and model theory in this last part of Section 3.5 may have left a rather motley impression. This is the effect of our decision to include in our language only those special predicates and functors which happened to be needed in the DRSs displayed earlier in the section. As we noted, a DRS language capable of representing, in a direct and natural way, the temporal information expressible in a language like English would require a much richer vocabulary, and would appear much less arbitrary than the one we have considered here.

¿From a methodological point of view, however, the language we have presented is not as arbitrary as it may seem. For the predicates and functors it contains exemplify between them a substantial part of the complications a model theory for a DRT-based language capable of presenting the various kinds of temporal information we find in natural language will have to deal with. The largest simple exception to this concerns the substantial range of concepts which natural languages employ for the description of metric concepts. There is only one pale reflection of this aspect of time in the language we represent here, viz. the predicate DAY. Its extension, we said, partitions the time into intervals. Intuitively these intervals are all of equal duration. But since in the language considered here DAY is the only metric notion, the metric aspect of its extension played no further part. For some of the issues connected with the model theoretic treatment of metric-related expressions of English see [Kamp and Schiehlen, 2002].

3.6 A First-Order DRT Calculus

When we ask whether a given conclusion that is presented in natural language follows from premises given in that same language it will quite often be the case that the conclusion depends for its interpretation in various ways on those premises. To take an extremely simple example, is the following argument valid:

(129) Peter ate a pizza and drank a glass of wine.
So, he ate something.

Here the second sentence does seem to follow from the first. But it does so only when we interpret *he* as anaphoric to *Peter*.

A natural way to capture the context-dependent notion of validity illustrated by this example is to construct a DRS K_{pr} for the premises of the argument and to then use this DRS K_{pr} as context for the construction of a DRS K_{con} for the putative conclusion. What we will typically get in this way is a pair consisting of (i) a proper DRS K_{pr} and (ii) a possibly improper DRS, but such that the merge of K_{pr} and K_{con} is again proper. Of this pair we can then ask whether the first DRS semantically entails the second, that is if any verifying embedding f of K_{pr} in any model $\mathcal{M}$ can be extended to a verifying embedding of K_{con} in $\mathcal{M}$. The following definition generalises this intuition. For technical reasons it allows for free discourse referents to occur in K_{pr} and K_{con}. Nevertheless K_{pr} and K_{con} must be *pure*, i.e. no discourse referent is allowed to be declared in two distinct DRSs, one subordinate to the other.

DEFINITION 40. For K and K' *pure* (but not necessarily *proper*) DRSs: K $\models_{\text{DRS}}$ K' holds iff for every model $\mathcal{M} = \langle$ U, $\Im$ $\rangle$ and embedding functions f and g such that $f \subseteq_{\text{U}_\text{K} \cup \text{FV(K)} \cup \text{FV(K')}} g$ such that $\langle f, g \rangle \models_{\mathcal{M}}$ K, there is a function h such that $g \subseteq_{\text{U}_{\text{K}'}} h$ such that $\langle g, h \rangle \models_{\mathcal{M}}$ K'.

In order to obtain a proof system for this notion of validity wrt. the first order DRS language presented in Section 3.1, there are two options. The first consists in mapping a proof argument $K_{pr} \vdash K_{con}$ into the formula of predicate logic that is the result of the translation of the DRS-condition $K_{pr} \Rightarrow K_{con}$ according to Def. (12) above and then employ any of the standard calculi developed for FOPL (viz. [Sundholm, 1986],[Sundholm, 2001]). The second option is to develop deduction rules that operate directly on DRT style proof representations $K_{pr} \vdash K_{con}$. [Koons, 1988; Sedogbo, 1988; Reinhard, 1989; Saurer, 1993; Reyle and Gabbay, 1994] and [Kamp and Reyle, 1991] provide a number of sound and complete proof systems of this type, obviating the detour through FOPL. In the following we will present the calculus presented in [Kamp and Reyle, 1991].

[Kamp and Reyle, 1991] represent premise conclusion pairs K_{pr} and K_{con} in the format used in [Kalish and Montague, 1964] and [Bonevac, 1987] with K_{con} occurring within a "Show-line" that is embedded within the premise DRS $K_{pr} = \langle\{x_1,...,x_n\},\{\gamma_1,...,\gamma_m\}\rangle$:

$$(130)\quad \boxed{\begin{array}{c} x_1 \;...\; x_n \\ \gamma_1 \\ \vdots \\ \gamma_m \\ \text{Show: } K_{con} \end{array}}$$

A proof is accomplished if the Show-line is cancelled, denoted by ~~Show:~~ K_{con}. Cancelling of a Show-line is achieved whenever one of the *rules of proof* has successfully been applied to it. Additional Show-lines may be added at any point in the derivation (provided only that merging the Show-line DRS with the DRS into which the Show-line is inserted would not result in an improper DRS). However, once a Show-line has been introduced it must be cancelled at a later time in order that the derivation counts as complete.

Rules of proof come in two types: *direct* and *indirect* rules of proof. Direct proofs do not involve any subproofs while indirect ones do. The system has one direct rule of proof RDP (Rule of Direct Proof) and two indirect rules of proof CP (Conditional Proof) and RAA (Reductio Ad Absurdum). In additon to the rules of proof there are *inference rules.* They apply to a DRS K_{pr} and extend it to a DRS K'_{pr} with $K_{pr} \subseteq K'_{pr}$. The system without disjunction and identity involves three inference rules DET (Detachment - also referred to as GMP (Generalized Modus Ponens)), DNE (Double Negation Elimination) and NEU (Non-Empty Universe). The full system with disjunction and identity features four additional inference rules MTP (Modus Tollendo Ponens), DI (Disjunction Introduction), SoI (Substitution of Identicals) and SI (Self-Identity). Soundness and completeness theorems relating $\models_{\mathrm{DRS}}$ and $\vdash_{\mathrm{DRS}}$ are proved in [Kamp and Reyle, 1991]. Moreover, the sublanguage involving "$\neg$" but without "$\Rightarrow$" and "$\vee$" requires only the rules of proof RDP and RAA and the inference rules DNE and NEU (as well as SOI and SI iff "$=$" is included as well). In each of these cases the system consisting of the mentioned inference rule and rules of proof is sound and complete for the model theory of Section 3.1.

SUMMARY 41. Architecture of a First-Order DRT Calculus

Rules of Proof			
DIRECT		INDIRECT	
RDP	Rule of Direct Proof	CP	Conditional Proof
		RAA	Reductio Ad Absurdum

Inference Rules	
DET	Detachment
DNE	Double Negation Elimination
NEU	Non-Empty Universe
MTP	Modus Tollendo Ponens
DI	Disjunction Introduction
SOI	Substitution of Identicals
SI	Self Identity

The Rule of Direct Proof (RDP) states that a DRS or DRS-condition Δ is proved if an alphabetic variant Δ' of Δ occurs as part of the DRS which contains the Show-line Show:Δ.

There are two notions in this description which have not yet been defined, "alphabetic variant" and "contains". The definition of "contains" is entirely straightforward.

DEFINITION 42. A DRS $\langle\, U_1, Con_1 \,\rangle$ *is contained in* a DRS $\langle\, U, Con \,\rangle$ iff $U_1 \subseteq U$ and $Con_1 \subseteq Con$.

The notion of alphabetic variant is most clearly defined for pure DRSs (see Definition 6). Since alphabetic variance enters into the formulation of several rules we will make things easy by restricting attention in this section to pure DRSs. (It can easily be verified that the changes produced by the application of the rules of the system preserve purity.)

DEFINITION 43. Let K and K′ be DRSs. Then K′ is an *alphabetic variant* of K iff there is a function f which maps the set BV(K) of bound discourse referents of K onto the bound discourse referents of K′ such that

(i) for each sub-DRS K″ of K $\mathrm{f}|_{\mathrm{U}_{\mathrm{K}''}}$ is one-to-one, and

(ii) K′ is the result of replacing for each $\mathrm{x} \in \mathrm{BV(K)}$ all occurrences of x in K by $f(\mathrm{x})$.

For the remainder of this section all DRSs will be pure.

DEFINITION 44. **Rule of Direct Proof (RDP):** if a DRS K contains a Show-line Show: Δ and if K contains Δ' where Δ' is an alphabetic variant of Δ, the Show-line may be cancelled.

Direct Proofs are proofs involving RDP and the inference rules only. The inference rule of Detachment (DET - also referred to as Generalised Modus Ponens) applies to DRS conditions of the form $\mathrm{K}_1 \Rightarrow \mathrm{K}_2$ in a DRS K. DET states that provided it is possible to homomorphically embed the antecedent K_1 into K we can add to K an alphabetic variant K_2' of the consequent K_2 such that

(i) the bound discourse referents of K_2' do not already occur in K, and

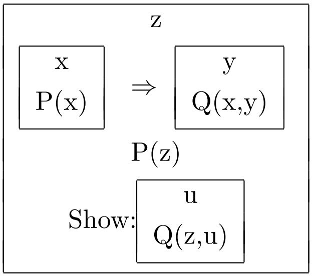
(ii) K_2' extends the homomorphic embedding f of K_1.

DEFINITION 45. **Detachment (DET) (Generalized Modus Ponens (GMP)):** Given a DRS K, if $\mathrm{K}_1 \Rightarrow \mathrm{K}_2 \in \mathrm{Con}_{\mathrm{K}}$ and if there is a homomorphic embedding $f(\mathrm{K}_1)$ into K, then we may add an alphabetic variant $g(\mathrm{K}_2)$ to K where $f \subseteq_{\mathrm{U}_{\mathrm{K}_2}} g$, $g \setminus f$ is one-to-one and g maps $\mathrm{U}_{\mathrm{K}_2}$ to a set of discourse referents that do not already occur in K.

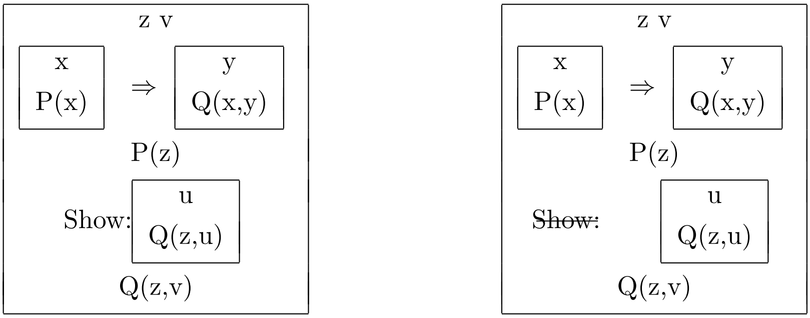
Definitions (44) and (45) can be illustrated with the following example. In order to show that

(131) $$\boxed{\begin{array}{c} \mathrm{z} \\ \boxed{\begin{array}{c} \mathrm{x} \\ \mathrm{P(x)} \end{array}} \Rightarrow \boxed{\begin{array}{c} \mathrm{y} \\ \mathrm{Q(x,y)} \end{array}} \\ \mathrm{P(z)} \end{array}} \quad \vdash_{\mathrm{DRS}} \quad \boxed{\begin{array}{c} \mathrm{u} \\ \mathrm{Q(z,u)} \end{array}}$$

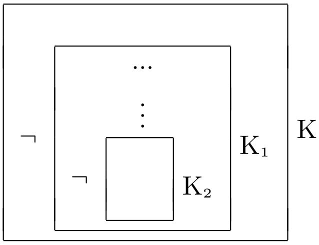
we add a Show-line with the conclusion to the premise in (131):

(132) 

Since the left-hand-side of the conditional DRS condition in (132) can be homomorphically embedded in the main DRS we can apply DET (45) and add an alphabetic variant of the right-hand-side, which extends the homomorphic embedding of the left-hand side, to the main DRS as shown on the left of (133). Then we apply RDP and cancel the Show-line yielding the proof structure shown on the right of (133), completing the proof of (131):

(133) 

The rule of Double Negation Elimination (DNE) applies to structures of the form

(134) 

In simple cases this amounts to the cancellation of two negation signs. In more complex cases where K_1 contains conditions other than $\neg$ K_2, DNE can be applied provided that K_1 - $\langle\emptyset, \{\neg K_2\}\rangle$ has a homomorphic embedding in K.

DEFINITION 46. **Double Negation Elimination (DNE):** if $\neg$ $K_1 \in \mathrm{Con}_K$ and $\neg$ $K_2 \in \mathrm{Con}_{K_1}$ and $f(K_1 - \langle\emptyset, \{\neg K_2\}\rangle)$ is a homomorphic embedding into K, then $g(K_2)$ may be added to K where $f \subseteq_{U_{K_2}} g$, $g - f$

is one-to-one and g maps the set of discourse referents $\mathrm{U}_{\mathrm{K}_2}$ to a set of discourse referents new to K.

The rule of Non-Empty Universe (NEU) states that we only consider models with non-empty universes. This means that we can always introduce discourse referents at the highest level of the DRS.

DEFINITION 47. **Non-Empty Universe (NEU):** if K is a DRS we may always add a new discourse referent to U_{K}.

Disjunction is treated in terms of two inference rules: Modus Tollendo Ponens (MTP) and Disjunction Introduction (DI). Modus Tollendo Ponens states that given a DRS with a disjunctive condition together with the negation of an alphabetic variant of one of the disjuncts we may add a disjunctive condition to the DRS which is like the original disjunction except that the disjunct corresponding to the negated condition is missing.

DEFINITION 48. **Modus Tollendo Ponens (MTP):** given an DRS K with a disjunctive condition of the form $\mathrm{K}_1 \vee \ldots \vee \mathrm{K}_{i-1} \vee \mathrm{K}_i \vee \mathrm{K}_{i+1} \vee \ldots \mathrm{K}_n$ and a condition of the form $\neg\mathrm{K}'_i$ where K'_i is an alphabetic variant of K_i we may add $\mathrm{K}_1 \vee \ldots \vee \mathrm{K}_{i-1} \vee \mathrm{K}_{i+1} \vee \ldots \mathrm{K}_n$ to K.

Disjunction Introduction permits us to introduce any disjunctive condition into a DRS if the DRS already contains one of the disjuncts.

DEFINITION 49. **Disjunction Introduction (DI):** if K_i is included in K then we may add $\mathrm{K}_1 \vee \ldots \vee \mathrm{K}_{i-1} \vee \mathrm{K}_i \vee \mathrm{K}_{i+1} \vee \ldots \mathrm{K}_n$ to K.

The proof system features two inference rules pertaining to identity: Substitution of Identicals (SoI) and Self-Identity (SI).

DEFINITION 50. **Substitution of Identicals (SoI):** if K contains conditions $\mathrm{x} = \mathrm{y}$ and γ where $\mathrm{x,y} \notin \mathrm{Decl}(\gamma)$, we may add a condition γ' to K where γ' results from γ by replacing one occurrence of x by y.

DEFINITION 51. **Self-Identity (SI):** if K is a DRS, then for any $\mathrm{x} \in \mathrm{U}_{\mathrm{K}}$ we may add $\mathrm{x} = \mathrm{x}$ to K.

As stated the inference rules apply at the level of the "main" DRS only. It can be shown, however, that the application of the inference rules can be *extended* to embedded DRSs and furthermore that every argument provable in the thus *extended proof system* is also provable in the old system.

The inference rules described above are based entirely on the premise DRS. Applying them extends the premise DRS until RDP can be applied. Proofs based on RDP and the inference rules are referred to as *direct* proofs. They do not involve any intermediate proofs and do not introduce any new temporary assumptions. In addition to direct proofs, the calculus features two rules of proof for *indirect* proofs involving sub-proofs and the introduction of temporary assumptions. These rules are the rule of Conditional Proof (CP) and Reductio ad Absurdum (RAA).

The rule of Conditional Proof is applied in proofs of DRS conditions of the form $K_1 \Rightarrow K_2$ in a premise DRS K. CP introduces a sub-proof which, on the assumption that an alphabetic variant of K_1 holds, tries to derive a variant of K_2. The sub-proof may make use of what is asserted in the premise DRS K. If the sub-proof is successful, $K_1 \Rightarrow K_2$ is established and the sub-proof and the temporary assumption are discarded.

DEFINITION 52. **Conditional Proof (CP):** if Con_K in a premise DRS K contains a Show-line Show: $K_1 \Rightarrow K_2$, we may introduce a sub-proof

$$\mathrm{K} \quad \| \quad \begin{array}{|c|}\hline \mathrm{K}'_1 \\ \text{Show:}\mathrm{K}'_2 \\ \hline\end{array}$$

where K'_1 and K'_2 are alphabetic variants of K_1 and K_2, respectively. When the Show-line in the sub-proof is cancelled, the Show-line Show: $K_1 \Rightarrow K_2$ in the premise DRS K may be cancelled as well.

Suppose we want to show

$$(135) \quad \begin{array}{|c|}\hline \begin{array}{|c|}\hline \mathrm{x} \\ \mathrm{P(x)} \\ \hline\end{array} \Rightarrow \begin{array}{|c|}\hline \mathrm{Q(x)} \\ \hline\end{array} \\ \begin{array}{|c|}\hline \mathrm{y} \\ \mathrm{Q(y)} \\ \hline\end{array} \Rightarrow \begin{array}{|c|}\hline \mathrm{R(y)} \\ \hline\end{array} \\ \hline\end{array} \quad \vdash_{\mathrm{DRS}} \quad \begin{array}{|c|}\hline \begin{array}{|c|}\hline \mathrm{z} \\ \mathrm{P(z)} \\ \hline\end{array} \Rightarrow \begin{array}{|c|}\hline \mathrm{R(z)} \\ \hline\end{array} \\ \hline\end{array}$$

We add the conclusion in a Show-line to the premise DRS and apply CP.

$$(136) \quad \begin{array}{|c|}\hline \begin{array}{|c|}\hline \mathrm{x} \\ \mathrm{P(x)} \\ \hline\end{array} \Rightarrow \begin{array}{|c|}\hline \mathrm{Q(x)} \\ \hline\end{array} \\ \begin{array}{|c|}\hline \mathrm{y} \\ \mathrm{Q(y)} \\ \hline\end{array} \Rightarrow \begin{array}{|c|}\hline \mathrm{R(y)} \\ \hline\end{array} \\ \text{Show: } \begin{array}{|c|}\hline \mathrm{z} \\ \mathrm{P(z)} \\ \hline\end{array} \Rightarrow \begin{array}{|c|}\hline \mathrm{R(z)} \\ \hline\end{array} \\ \hline\end{array} \quad \| \quad \begin{array}{|c|}\hline \mathrm{z} \\ \mathrm{P(z)} \\ \text{Show:} \begin{array}{|c|}\hline \mathrm{R(z)} \\ \hline\end{array} \\ \hline\end{array}$$

Note well: In connection with proving the Show-line to the right of || the entire DRS to the left of || is available as premise (with the exception of course of the Show-line, or Show-lines it contains). Put differently, the premise DRS for the Show-line on the right is the merge of the DRSs to the left and right of || without their respective Show-lines. In this regard the architecture of the present system is like that of any other natural deduction system.

Now we can apply DET twice: from $\langle$ {z}, {P(z)} $\rangle$ in the CP sub-derivation and the first condition in the premise DRS we get Q(z) and from $\langle$ {z}, {Q(z)} $\rangle$ together with the second condition in the premise DRS we get R(z).

(137)
$$\boxed{\begin{array}{c} \boxed{\begin{array}{c} x \\ P(x) \end{array}} \Rightarrow \boxed{\begin{array}{c} \\ Q(x) \end{array}} \\ \boxed{\begin{array}{c} y \\ Q(y) \end{array}} \Rightarrow \boxed{\begin{array}{c} \\ R(y) \end{array}} \\ \text{Show: } \boxed{\begin{array}{c} z \\ P(z) \end{array}} \Rightarrow \boxed{\begin{array}{c} \\ R(z) \end{array}} \end{array}} \quad \| \quad \boxed{\begin{array}{c} z \\ P(z) \\ \text{Show: } \boxed{\begin{array}{c} \\ R(z) \end{array}} \\ Q(z) \\ R(z) \end{array}}$$

Now RDP may be applied to the CP sub-derivation cancelling the Show-line Show: $\boxed{\begin{array}{c} \\ R(z) \end{array}}$. According to the CP rule we may also cancel the Show-line in the premise DRS, completing the proof of (135).

The final rule of proof, Reductio ad Absurdum, also opens up a new sub-derivation in which we try to show that the assumption $\neg$ K_1' where K_1' is an alphabetic variant of K_1 and K_1 is a goal in a Show-line in the premise DRS, leads to an explicit contradiction thus establishing K_1. Here by an explicit contradiction we mean the following. A DRS K contains an explicit contradiction iff there is a DRS K' such that

(i) $\neg$ $K_1' \in Con_K$, and

(ii) K contains an alphabetic variant of K'.

We use $\perp$ to represent arbitrary contradictions of this kind. Thus "Show:$\perp$" can be cancelled when the DRS contianing this Show-line also contains such a combination of $\neg$ K' and a variant of K' (for any K' whatever).

DEFINITION 53. **Reductio ad Absurdum (RAA):** if Con_K of some premise DRS K contains a Show-line Show: K_1 we may introduce a sub-proof

$$K \quad \| \quad \boxed{\begin{array}{c} \neg\ K_1' \\ \text{Show:}\perp \end{array}}$$

When the Show-line in the sub-proof is cancelled, the Show-line "Show: K_1" in the premise DRS K may be cancelled as well.

One place where RAA is needed is in the proof of the principle of Modus Tollens, which in the present system is a derived rather than a primitive rule. The following example shows one variant of this principle.

$$(138)\quad \boxed{\begin{array}{c} \boxed{\begin{array}{c} \mathrm{x\ y} \\ \mathrm{P(x,y)} \end{array}} \Rightarrow \boxed{\begin{array}{c} \mathrm{u} \\ \mathrm{Q(u,x)} \end{array}} \\ \neg \boxed{\begin{array}{c} \mathrm{v\ t} \\ \mathrm{Q(v,t)} \end{array}} \end{array}} \quad \vdash_{\mathrm{DRS}} \quad \boxed{\neg \boxed{\begin{array}{c} \mathrm{x\ y} \\ \mathrm{P(x,y)} \end{array}}}$$

We add the conclusion in a Show-line and apply RAA and DNE.

$$(139)\quad \boxed{\begin{array}{c} \text{(i)}\ \boxed{\begin{array}{c} \mathrm{x\ y} \\ \mathrm{P(x,y)} \end{array}} \Rightarrow \boxed{\begin{array}{c} \mathrm{u} \\ \mathrm{Q(u,x)} \end{array}} \\ \text{(ii)}\ \neg \boxed{\begin{array}{c} \mathrm{v\ t} \\ \mathrm{Q(v,t)} \end{array}} \\ \text{Show: } \neg \boxed{\begin{array}{c} \mathrm{w\ z} \\ \mathrm{P(w,z)} \end{array}} \end{array}} \quad \| \quad \boxed{\begin{array}{c} \mathrm{w'\ z'} \\ \mathrm{P(w',z')} \\ \text{Show: } \bot \end{array}}$$

By applying DET on (i) and the DRS on the right of $\|$ we obtain:

$$(140)\quad \boxed{\begin{array}{c} \text{(i)}\ \boxed{\begin{array}{c} \mathrm{x\ y} \\ \mathrm{P(x,y)} \end{array}} \Rightarrow \boxed{\begin{array}{c} \mathrm{u} \\ \mathrm{Q(u,x)} \end{array}} \\ \text{(ii)}\ \neg \boxed{\begin{array}{c} \mathrm{v} \\ \mathrm{Q(v,x)} \end{array}} \\ \text{Show: } \neg \boxed{\begin{array}{c} \mathrm{w\ z} \\ \mathrm{P(w,z)} \end{array}} \end{array}} \quad \| \quad \boxed{\begin{array}{c} \mathrm{w'\ z'\ u'} \\ \mathrm{P(w',z')} \\ \text{Show: } \bot \\ \mathrm{Q(u',w')} \end{array}}$$

We now have $\neg \boxed{\begin{array}{c} \mathrm{v\ t} \\ \mathrm{Q(v,t)} \end{array}}$ as a condition in the DRS to the left of $\|$ and a variant $\boxed{\begin{array}{c} \mathrm{u'\ w'} \\ \mathrm{Q(u',w')} \end{array}}$ of the DRS in the scope of the negation contained in the (extended) DRS to the right of $\|$. This establishes the contradiction and we can cancel the Show-line "Show: $\bot$" on the right and with it the Show-line in the DRS on the left, completing the proof.

Note also that the DRS version of the argument $\neg(A \wedge \neg B), A \vdash B$ (a version of Modus Ponens) can be proved by a simple application of RAA. For example, the Show-line in (141.a) can be derived by adding the environment for an application of RAA to the right of it as shown in (141.b).

(141) a. $\begin{array}{|c|}\hline z \\ P(z) \\ \neg \begin{array}{|c|}\hline x \\ P(x) \\ \neg\begin{array}{|c|}\hline y \\ Q(x,y) \\ \hline\end{array} \\ \hline\end{array} \\ \text{Show: } \begin{array}{|c|}\hline v \\ Q(z,v) \\ \hline\end{array} \\ \hline\end{array}$ b. $\begin{array}{|c|}\hline z \\ P(z) \\ \neg \begin{array}{|c|}\hline x \\ P(x) \\ \neg\begin{array}{|c|}\hline y \\ Q(x,y) \\ \hline\end{array} \\ \hline\end{array} \\ \text{Show: } \begin{array}{|c|}\hline v \\ Q(z,v) \\ \hline\end{array} \\ \hline\end{array} \; \| \; \begin{array}{|c|}\hline \neg\begin{array}{|c|}\hline w \\ Q(z,w) \\ \hline\end{array} \\ \text{Show:}\bot \\ \hline\end{array}$

It now suffices to observe that if we add the new assumption $\neg \begin{array}{|c|}\hline y \\ Q(z,y) \\ \hline\end{array}$ to the DRS on the left, then this DRS will contain an alphabetic variant of the condition $\neg \begin{array}{|c|}\hline x \\ P(x) \\ \neg\begin{array}{|c|}\hline y \\ Q(x,y) \\ \hline\end{array} \\ \hline\end{array}$ which belongs to its condition set. So the Show-line "Show: $\bot$" can be cancelled and with it the Show-line on the left.

We noted earlier that the fragment of the DRS language of Section 3.1 in which only complex conditions are of the form $\neg$ K has the same expressive power as the full language. For this sublanguage the present system reduces to one consisting of the rules RDP, NEU, DNE and RAA. This system is sound and complete for the given fragment, just as the full system is sound and complete for the full language of Section 3.1.

One of the features of DRS languages, we stressed in Section 3.3, is that they do not separate sentential and quantificational aspects in the manner familiar from standard predicate logic. This feature of the syntax of DRS languages has its reflection in the rules of the deduction system we have presented. It is manifest in every rule which involves matching of alphabetic variants. This feature we have seen, is particularly prominent in applications of DET and RAA, and indeed it is only because RAA is stated as applicable in cases of a contradiction between alphabetic variants that DNE, NEU, RDP and RAA suffice for the fragment which is without "$\Rightarrow$", "$\vee$" and "=". But it is indispensible also in the presence of other rules such as DET and CP. Without this flexibility in the application of RAA the DR-theoretical equivalent of $\neg(\mathrm{A} \wedge \neg \mathrm{B}), \mathrm{A} \vdash \mathrm{B}$ cannot be derived even when all the other rules are available. In short, matching of alphabetic variants in the application of deduction rules is the proof theoretic mirror of the structural binding of discourse referents (through membership in a certain DRS-universe) which is perhaps the most distinctive feature of the DR-theoretical representation format.

Given that the present deduction system clearly reflects this feature of the DRS language to which the system applies, it would appear to be of interest (i) to extend it with rules that equally mirror this feature of DRT for the extended languages we have discussed in Sections 3.3–3.5 (Of course for the non-axiomatisable extensions, such as that of 3.4, such a coverage could only be partial.); and (ii) to explore the possibilities of implementations of such proof procedures. To our knowledge neither of these tasks has thus far been persued in good depth.

4 PRESUPPOSITION

4.1 Introduction

Dynamic Semantics is ideally suited to the analysis of presupposition. This is true of all versions of it, including the two that come first historically, File Change Semantics (FCS, see [Heim, 1982]) and DRT. As we have seen in the previous sections for DRT, a central rationale for these theories was to give a context-based account of pronominal anaphora. In this section we will see how such an account can be extended to a context-based account of presupposition.

To deal with cases of transsentential anaphora one needs a formally precise notion of context. All Dynamic theories provide such a notion, a notion of "discourse" context which evolves as the discourse proceeds, with each new sentence meaning its own contribution to it. Each sentence is to be interpreted in the current discourse context, and thus in the light of what its predecessors have contributed to it. The notion of discourse context can be refined, moreover, so that it can change even in the course of a single sentence, with some parts of the sentence contributing to the context serving the interpretation of some part. We already saw that along these lines it is possible to develop a uniform account of transsentential and sentence-internal anaphora.

The Dynamic concept of a discourse context which changes not only between sentences but also sentence-internally is crucial not only for an account of anaphora but also of presupposition. In particular it is essential for dealing with the so-called Projection Problem. Sometimes a presupposition that is generated within some part of a logically complex sentence is perceived as presupposition of the entire sentence — the presuppositon "projects" — and sometimes it seems to have disappeared when one considers the sentence as a whole — the presupposition does not "project". The basic strategy that the Dynamic approach offers for explaining this difference is surprisingly simple: A presupposition doesn't project if it is justified by its "local" context, i.e. on the basis of contextual information that is entirely sentence-internal. For in that case its justification has no further need for information from the "global", or sentence-external, context; so, as far as this presuppositon is concerned any global context whatever would be a suitable context in which the sentence could, as far as presuppositions are concerned, be properly used.

The parallel that is suggested by this gloss on presupposition projection is too obvious to overlook: When a pronoun has a sentence-internal antecedent — i.e. when it finds an antecedent in its local context — it is no obstacle to interpreting the sentence as one which expresses a proposition on its own, and no further contextual information is required. Likewise for a locally justified presupposition. Only when pronoun or presupposition cannot be accounted for on the basis of sentence-internal information alone does their presence turn into a constraint on the global context — to provide an antecedent for the pronoun or to justify (or assist in justifying) the presupposition.

Indeed, it was not long after FCS and DRT were first proposed that Heim formulated an account of presupposition which extends the Dynamic approach to anaphora to presuppositional phenomena, and most notably to the Projection Problem [Heim, 1983]. But it wasn't until the very end of the eighties that the central ideas of her proposal were pushed further. At that point a number of people proposed an even more tightly unified account of presupposition and anaphora. (See [Geurts and van der Sandt, 1999; van der Sandt, 1992; Zeevat, 1992]). In these proposals anaphoric expressions (and especially pronouns) are treated as "presupposition triggers", on a par with the presupposition triggers which in the theory of presupposition had long been recognised as such: definite descriptions, factive verbs like *regret*, *be surprised*, etc, aspectual verbs like *stop* or *continue*, particles like *again* or *too*, cleft-constructions, and so on (as many readers will surely know, the complete list is much, much longer). The presupposition triggered by an anaphoric expression is that an antecedent for it can be found in the context. The proposals that have just been mentioned are all formulated within the framework of DRT.

One consequence of such a unified treatment of presupposition and anaphora is that anaphoric expressions impose, just like other presupposition triggers, constraints in context. At the same time such a treatment highlights the "anaphoric" dimension of arbitrary presuppositions: Not only pronominal "presuppositions" act as pointers to information provided earlier, this is a feature of presuppositions in general; all presuppositions are "anaphoric" in the sense of linking the sentence or sentence part in which they originate with the relevant part of the context that serves as background for the interpretation of that sentence or sentence part. In this way, i.e. by linking a sentence or sentence constituent to those parts of the context where the required information is found, presuppositions foster and consolidate discourse coherence. As we will see below, this cohesion-creating effect of presuppositions is closely connected with presupposition *accommodation*, i.e. with the adaptation of an initially insufficient context in such a way that the given presuppositions can be seen to be justified after all by the adjusted context.

Our exposition will proceed as follows. In Section 4.2 we first give some elementary illustrations of how the present account of presupposition works, using examples which are taken from [van der Sandt, 1992] (modulo some trivial alternations). The notation we use differs cosmetically from the one found in Van Der Sandt's paper. More importantly, our treatment of definite descriptions differs from his, as well as from his and our treatment of anaphoric pronouns. This is an issue to which we devote a somewhat longer discussion, motivated by the consideration that the logical and philsophical tradition has for the most part treated descriptions and pronouns as separated by a major divide, with pronouns the paradigmatic variables of natural language and definite descriptions the prototypical presupposition triggers. Following this some further variants on the pattern of these examples are discussed. These variants are chosen in order to illustrate local presupposition justification, as the source of non-projection.

Section 4.2 ends with a few examples which bring out some of the complexities that arise when other kinds of presuppositions are taken into account besides those on which Van Der Sandt's paper focusses and to which we limit ourselves in the first two parts of this section. A further aim of this section is to reveal some of the intricate interactions that are often found between presuppositions connected with different presupposition triggers occurring within one and the same sentence. Section 4.3 presents the syntax for the DRT formalism in which the preliminary representations of the present account are expressed, and a model-theoretic semantics to go with it. As part of this we define the notions of global and local context, as well as one way of distinguishing between "anaphoric" and "non-anaphoric" presuppositions. Section 4.4 is devoted to presupposition resolution and accommodation and Section 4.5 to the principles according to which preliminary representations are constructed from syntactic trees.

4.2 Examples

Pronouns and Definite Descriptions in Simple Sentences

We begin by looking in some detail at the following examples (cf [van der Sandt, 1992]).

(142) a. Walter has a rabbit and a guinea pig. The rabbit is white.

b. Walter has a rabbit and a guinea pig. His rabbit is white.

c. Walter has a rabbit. It is white.

We start with (142.a). We assume that processing of the first sentence yields the DRS (143), and that this DRS represents the context within which the second sentence, *It is white.*, is to be interpreted.[33]

(143)

w y z
Walter(w) rabbit(y) guinea pig(z) have (w,y) have (w,z)

The preliminary representation of the second sentence contains a presupposition that is triggered by the definite description *the rabbit.* What form should the representation of this presupposition take? This question leads us directly to one of those central issues in the theory of presuppositions which the Dynamic approach has brought into sharper focus. Definite descriptions have been considered the prototypical cases of presupposition-triggering, since the time where the notion of presupposition was recognised as important to the theory of meaning and logic. In fact, it was they who gave rise to this issue in the first place. Frege, one of the two fathers of modern formal logic,[34] noted that the referential function of singular definite descriptions of the form *the N* is compromised by failure of either the associated existence condition — *there is at least one N* – or the associated uniqueness condition — *there is at most one N.* So he saw the conjunction of these two conditions as the presupposition that must be satisfied in order that the description can perform the function for which it is intended: refer to the unique x such that $N(x)$. If the presupposition is not satisfied, then, Frege thought, any sentence containing the description will

[33]In dealing with the examples in (142) we revert to the mode of representation in which temporal relations are ignored. We will return to representations which take temporal reference into account later on (starting with example (163)).

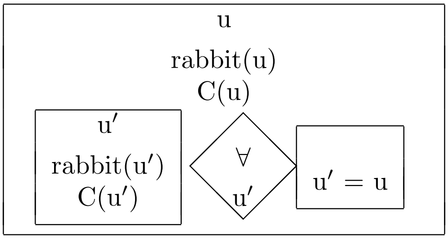
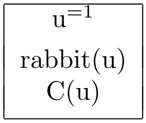
[34]We take it as established that the predicate calculus, the fundament on which all modern logic rests, was invented — or, if you prefer, discovered — independently by Frege and Peirce.

fail to have a proper truth value, with unforseeable consequences for the logic of formal systems into which definite descriptions are admitted. The proposals to which Frege's worries about presupposition gave rise during the following 75 years seem to have been concerned almost exclusively with definite descriptions, from Russell's Theory of Descriptions [Russell, 1905] to Strawson's revindication of the Fregean perspective [Strawson, 1950; Strawson, 1964], and the literature that arose out of the debate provoked by Strawson's 1950 publication in the course of the years following it.

It was not until the late sixties that presupposition became an active concern within linguistics. One important effect of this was that presupposition came to be seen as a much more general phenomenon, of which the presuppositions of definite descriptions are only one among many different manisfestations. But even since that time the presuppositions of definite descriptions have retained much of their paradigmatic status.

As said, the logicians of the end of the 19-th and the first half of the 20-th century took the presupposition of a singular definite description *the N* to be the proposition that there exists a unique individual satisfying the predicate *N*. It cannot but have been clear from the start that the definite descriptions used in ordinary conversation hardly ever satisfy this proposition when it is taken literally. (142.a) is a case in point. Noone who hears (142.a) will take it to imply that there is only one rabbit in the entire universe. Insofar as the uniqueness requirement applies to this case, it is only in the sense that the satisfier of the predicate *rabbit* is uniquely determined within the context in which the sentence containing the description (the second sentence of (142.a)) appears. This context can be seen as providing a restricted set of individuals, and it is only within this set that *rabbit* can be assumed to have a unique satisfier. In our example this condition is fulfilled when we take the context to be given by (143), and the context set as given by its DRS-universe $\{w,y,z\}$. For in light of the information which (143) makes available about the represented entities, it seems safe to conclude that only one of them is a rabbit.

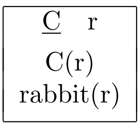
It follows that a plausible version of the existence-and-uniqueness presupposition for singular descriptions will have to allow for contextual restriction. We represent this restriction in the form of a predicate C (cf. [von Fintel, 1994]), ??). In particular, the representation of the existence-and-uniqueness presupposition for *the rabbit* in (142.a) takes the form given in (144.a).

(144) a.  b. 

We will abbreviate a DRS representing the existence-and-uniqueness presupposition for singular descriptions by superscribing =1 to the discourse referent representing the individual the singular description denotes. I.e. the DRS in (144.a) will be abbreviated by (144.b).

The contextual predicate C must, as the term "contextual" implies, be "recovered" from the context in which the description is used. Thus C imposes on the context a constraint which is reminiscent of those imposed by anaphoric pronouns: the predicate C is to be identified with this "antecedent" and the identification should fit the interpretation of the discourse as a whole — more specifically, it should enable the interpreter to see the contextualised existence-and-uniqueness presupposition as fulfilled.

This leads us to the conclusion that the existence-and-uniqueness presupposition of a definite description presupposition comes with a further, more "anaphoric" presupposition, to the effect that an antecedent must be found for the predicate variable C. We represent this latter presupposition in the form given in the DRS (145). (145) treats C as a discourse referent of higher type (that of a predicate of individuals). The only constraint on C, which is entailed by the role that it plays in the existence-and-uniqueness presupposition from which its presupposition derives, is that there must be at least one thing falling under C which satisfies the overt descriptive content of the description — i.e., in the case of our example, that there must be at least one thing in C's extension which is a rabbit. The underlining of C in the universe of (145) serves as indication that C is anaphoric, i.e. that the context must provide a suitable value for it.

(145) 

The classical view of the contribution that is made by a definite description to the proposition expressed by the sentence in which it occurs is as follows. On the assumption that the existence-and-uniqueness presuppositon of the description is satisfied, the proposition expressed is that some instance of the descriptive content – or, if one prefers, its unique instance; in case the presupposition is satisfied, the distinction doesn't matter — satisfies the predicate which the sentence asserts of the descriptive NP. Thus, in our example (142.a) the proposition expressed by the second sentence is that some rabbit (viz. the contextually unique one) is white.

We take it to be implied by this perspective of what proposition is expressed that in a case where the descriptive content is reinforced by a contextual predicate C, this additional predication also becomes part of the content of the proposition. So in particular, the proposition expressed by the second sentence of (142.a) is that some rabbit with the property C is white. Consequently the non-presuppositional part of the preliminary representation must be to the effect that there is something which is a rabbit, satisfies C and is white:

$$(146)\quad \begin{array}{|c|}\hline \mathrm{v} \\ \\ \mathrm{rabbit(v)} \\ \mathrm{C(v)} \\ \mathrm{white(v)} \\ \hline \end{array}$$

We represent presuppositions as left-adjoined to those parts of the preliminary sentence representation which represent the parts of the sentence which contain their trigger. Moreover, presuppositions which are generated by other presuppositions are left-adjoined to the representations of those. In the case of the second sentence of (142.a) this means that the existence-and-uniqueness presupposition gets adjoined to the representation of the sentence as a whole, while the anaphoric presuppositon concerning C gets left-adjoined to the existence-and-uniqueness presupposition. Thus we arrive at the preliminary representation in (6).[35]

$$(147)\quad \left\langle \left\{ \left\langle \left\{ \begin{array}{|c|}\hline \underline{\mathrm{C}}\ \mathrm{r} \\ \\ \mathrm{C(r)} \\ \mathrm{rabbit(r)} \\ \hline \end{array} \right\}, \begin{array}{|c|}\hline \mathrm{u}^{=1} \\ \\ \mathrm{rabbit(u)} \\ \mathrm{C(u)} \\ \hline \end{array} \right\rangle \right\}, \begin{array}{|c|}\hline \mathrm{v} \\ \\ \mathrm{rabbit(v)} \\ \mathrm{C(v)} \\ \mathrm{white(v)} \\ \hline \end{array} \right\rangle$$

The final representation of the discourse (142.a) is obtained by combining (147) with the context DRS (143). This combination involves justification of the two presuppositions of (147), the existence-and-uniqueness presupposition and the "anaphoric" presupposition concerning C that is adjoined to it. Resolution of the latter can, we have seen, take the form of identifying the extension of C with the the DRS-universe {w,y,z} of the context DRS. (Note that this resolution can be seen to satisfy the constraints of the C-presupposition, since the context (143) carries the information that one of the three represented individuals, that represented by y, is a rabbit.)

[35]The presence of the curly brackets is explained as follows: In general what gets adjoined to a given part of the representation is not a single presupposition, but a set of them. In (147) both sets are singletons. (When this is the case, the brackets may be omitted without risk of ambiguity.)

The effect of this is shown in (148). The C-presupposition has been eliminated now that the identification of C with the predicate "$\in\{w,y,z\}$" has led to its satisfaction.

$$(148)\quad \left\langle \left\{ \boxed{\begin{array}{c} u^{=1} \\[6pt] \text{rabbit}(u) \\ v\in\{w,y,z\} \end{array}} \right\}, \boxed{\begin{array}{c} v \\[6pt] \text{rabbit}(v) \\ \text{white}(v) \end{array}} \right\rangle$$

The remaining presupposition can now be seen as satisfied by the context DRS (143). But note that to "see" this we must rely on certain assumptions about the world (i.e. on "world knowledge"): (i) The assumption that an individual who owns rabbits and guinea pigs may be assumed to be a person; and (ii) the general knowledge that neither persons nor guinea pigs are rabbits. Such considerations very often enter into the justification of presuppositions. In further examples we will take this world knowledge-related aspect of presupposition justification for granted. It is important, however, to keep in mind how common it is for world knowledge to play some role in presupposition justification.

Given that the presupposition of (148) is justified in (143), it can be discarded as well, and the non-presuppositional part of (148) merged with (143). The result is the DRS (149).[36]

$$(149)\quad \boxed{\begin{array}{c} w\ y\ z\ v \\[6pt] \text{Walter}(w) \\ \text{rabbit}(y)\quad ;\quad \text{guinea pig}(z)\quad ;\quad \text{rabbit}(v) \\ \text{have }(w,y)\quad ;\quad \text{have }(w,z) \\ v\in\{w,y,z\} \\ \text{white}(v) \end{array}}$$

Now consider sentence (142.b). Once more we assume that (143) is the context representation in relation to which the second sentence of the discourse is interpreted. This time the presupposition-triggering subject, the definite description *his rabbit*, contains another definite NP, the pronoun *his*. Since we are aiming for a unified analysis, in which anaphoric expressions are treated as presupposition triggers too, our preliminary representation should also contain a presupposition associated with *his*. We will assume here without further argument that the possessive pronoun *his* can be analysed as decomposable into (i) the masculine singular pronoun (other realisations of which are the forms *he* and *him*) and (ii) a relation expressed

[36] Whenever atomic DRS-conditions are listed in one line we will separate them with ";". Note that this use of ";" is not dynamic conjunction — although there would be no truth-conditional difference for the case of atomic conditions. It is only a representational means to separate the elements in the condition set of a DRS.

by the genitival ending *'s*, which we will take to express a relation of possession between the referent of the pronoun and that of the NP containing it (here: the NP *his rabbit*). For simplicity we will represent this relation as "have(-,-)", just as we have been representing the verb *have* occurring in the first sentence of (142.a).

But how should we represent the presupposition triggered by the pronoun? We proceed in much the same way we did in connection with the contextual predicate C: The presupposition presents a discourse referent for the anaphoric element, here x, as requiring an antecedent. Since this is the purport of the presupposition, x appears with underlining (like C in its presupposition in (147)). The choice of x's antecedent is constrained by some information which the pronoun itself contributes. We make the simplifying assumption that the use of this information, carried by the English third person singular masculine pronoun, is that its referent must be a male person. Note that the underlined discourse referents of anaphoric presuppositions recur in the adjunction sites of these presuppositions. Non-underlined discourse referents, such as u in the existence-and-uniqueness presuppositions of (147) and (148) or r in the presupposition for C in (147), do not.

In (150) below the presupposition associated with the definite description is represented in the same way as before, viz as an existence-and-uniqueness presupposition involving a potential contextual restriction C. The pronoun *his* is part of the definite description *his rabbit* which gives rise to this presupposition. So the presupposition triggered by the pronoun arises in the process of interpreting the content of the description: it is a presupposition which must be resolved in order to determine what this content is. Note in this connection that the discourse referent x, which in the pronoun presupposition plays the role of anaphoric discourse referent in search of an antecedent, also occurs in the specification of the descriptive content of the existence-and-uniqueness presupposition of the definite description (as well as, by implication, in the representation of the proposition expressed by the sentence). Moreover, since the descriptive content is contextually restricted by C, the resolution of the pronoun is also relevant to the resolution of C. So we assume that the pronoun presupposition is left-adjoined to the complex presupposition. With these assumptions we arrive at the preliminary representation in (150).

(150) $$\left\langle \left\{ \left\langle \left\{ \boxed{\begin{array}{c} \underline{\mathrm{x}} \\ \\ \mathrm{male(x)} \\ \mathrm{pers(x)} \end{array}} \right\}, \left\langle \left\{ \boxed{\begin{array}{c} \underline{\mathrm{C}}\ \mathrm{r} \\ \\ \mathrm{C(r)} \\ \mathrm{rabbit(r)} \\ \mathrm{have(x,r)} \end{array}} \right\}, \boxed{\begin{array}{c} \mathrm{u}^{=1} \\ \\ \mathrm{rabbit(u)} \\ \mathrm{C(u)} \\ \mathrm{have(x,u)} \end{array}} \right\rangle \right\rangle \right\}, \boxed{\begin{array}{c} \mathrm{v} \\ \\ \mathrm{rabbit(v)} \\ \mathrm{C(v)} \\ \mathrm{have(x,v)} \\ \mathrm{white(v)} \end{array}} \right\rangle$$

Justification of the presuppositions of (150) within the context DRS (143) proceeds much as before. We now have one more presupposition to deal with, viz. the anaphoric presupposition triggered by *his*. The obvious resolution of this presupposition is that which identifies x with w. Once again the resolution of C to "$\in$ {w,y,z}" satisfies the constraints of the C-presupposition itself and guarantees justification of the existence-and-uniqueness presupposition (given the same uncontroversial bits of world knowledge). The resulting representation (151) for (142.b) is nearly the same as that for (142.a) and represents the same truth conditions.

$$
(151)\quad \begin{array}{|c|}
\hline
\text{w y z x v} \\
\text{Walter(w)} \\
\text{rabbit(y)} \ \ ; \ \ \text{guinea pig(z)} \\
\text{have (w,y)} \ \ ; \ \ \text{have (w,z)} \\
\text{x} = \text{w} \\
\text{rabbit(v)} \\
\\
\text{v}\in\{\text{w,y,z}\} \\
\text{have (x,y)} \\
\text{white(v)} \\
\hline
\end{array}
$$

Our third example, (142.c), differs from (142.a) and (142.b) in two respects: (i) the first sentence only introduces a rabbit into the discourse, but no guinea pig; and (ii) the subject NP of the second sentence is not a definite description but the pronoun *it*. What has been said in connection with the previous two examples largely determines the way in which we are to deal with this one. The context representation for the second sentence (that is, the representation for the first sentence of (142.c)) is the one given in (152). The preliminary representation, in (153), has only one presupposition, triggered by *it*. Just as we did in connection with *his* im (153.b), we simplify the constraints which it imposes on what sort of entity its referent can be, assuming simply that its referent must be a non-person. Note also that, like we saw in (150) for the discourse referent introduced by *his*, the distinguished discourse referent u of the anaphoric presuppositon recurs in the non-presuppositional part; lastly, the result of combining (152) and (153) yields, via the only conceptually admissible resolution of u (the one which identifies u with y), the DRS in (154).

$$
(152)\quad \begin{array}{|c|}
\hline
\text{w y} \\
\text{Walter(w)} \\
\text{rabbit(y)} \\
\text{have (w,y)} \\
\hline
\end{array}
$$

$$
(153)\quad \left\langle \left\{ \begin{array}{|c|}
\hline
\underline{\text{u}} \\
\text{non-pers(u)} \\
\hline
\end{array} \right\} , \begin{array}{|c|}
\hline
\text{white(u)} \\
\hline
\end{array} \right\rangle
$$

(154) $$\boxed{\begin{array}{c} \text{w y u} \\ \text{Walter(w)} \quad ; \quad \text{rabbit(y)} \quad ; \quad \text{have (w,y)} \\ \text{u} = \text{y} \\ \text{white(u)} \end{array}}$$

In the treatment of the examples (142.a-c) we have stuck as closely as possible to the traditional distinction between (a) definite descriptions as expressions whose denotation presupposes existence and uniqueness of descriptive content and (b) pronouns as anaphoric expressions, whose intepretation requires that they must be found an antecedent. Since, as we have seen, the definite descriptions of our examples cannot be analysed in this classical manner unless we allow for contextual restriction of their descriptive content, the two analysis strategies do not appear as radically different as they seem according to the logical picture that emerges from the by now "classical" literature in the philosophy of language, including the writings of Frege, Strawson and Quine, according to which a definite description is a singular term the use of which is subject to the truth of a certain presupposed proposition, while anaphoric pronouns are seen as the "variables of natural language".

Nevertheless, it might be thought that we haven't pushed the unified treatment of pronouns and definite descriptions far enough. In fact, many current analyses of definite descriptions treat them (or at any rate treat many of them) much more on a par with pronouns than we have done here.[37] In these analyses definite descriptions introduce, like pronouns, anaphoric discourse referents, while their descriptive content is treated as a restriction that must be satisfied by the antecedent for this discourse referent. For example, the definite description *the rabbit* in the second sentence of (142.a) gives rise, on such an analysis, to the anaphoric presupposition shown in (155).

(155) $$\left\langle \left\{ \boxed{\begin{array}{c} \underline{\text{u}} \\ \text{rabbit(u)} \end{array}} \right\}, \boxed{\begin{array}{c} \\ \text{white(u)} \end{array}} \right\rangle$$

On the analysis of definite descriptions which (155) exemplifies they are anaphoric NPs, which differ from pronouns only in that they are capable of providing more specific descriptive content. Favouring such a closely parallel treatment of pronouns and descriptions is the following consideration. Compare (142.c), in which the use of the pronoun *it* is coherent and unambiguous, with (156), in which it is not.

(156) *Walter has a rabbit and a guinea pig. It is white.

[37]An example is the paper [van der Sandt, 1992] itself, which has been the major inspiration for the theory sketched in this section.

The incoherence of *it* in (156) derives, it would appear, from its inability to distinguish between the two non-persons represented in the context, the rabbit and the guinea pig. (A description which doesn't discriminate between these two, such as, say, *the rodent* or *the furry creature*, would do just as poorly.) But *the rabbit* does fine, given that we all know that guinea pigs aren't rodents and that Walter, who "has" a rabbit, is therefore presumably not a rabbit himself.[38] It might seem from these considerations, that anaphoric pronouns and definite descriptions differ only in their descriptive content.

However, there are also considerations on the other side, which speak against such a rapprochement between our analysis of descriptions and pronouns. Arguments to this effect can be found in many places in the philosophical and semantic literature. Here we mention only one, which has to do with "bridging". Compare the pair of discourses in (157).

(157) a. Bill is a donkey owner. The donkey is not happy.

b. Bill is a donkey owner. ? It is not happy.

The interpretation of the donkey in (157.a) can be justified as follows: The first sentence entails that there are one or more donkeys that Bill owns. So this information can be regarded as part of the context in which the description has to find its reference. In order to justify the singular definite NP we have to "accommodate" the assumption that Bill's donkey ownership involves a single donkey only, but in the interpretation of (157.a) this does not appear to pose a problem. Consider now (157.b). Here too the content of the first sentence allows us to extend the context unverse from the set consisting just of Bill (the only individual explicitly mentioned) to one which contains in addition the donkey or donkeys he owns. If we suppose that the only distinction between *the donkey* and *it* concerns the desriptive contents of their respective presuppositions — that of *the donkey* is to the effect that its antecedent satisfies the predicate "donkey" and that of the pronoun that it satisfies the predicate "non-person" — then the fact that (157.a) is fine but (157.b) is not, becomes a mystery. For the predicate "non-person" is all we need to distinguish the donkey or donkeys owned by Bill from Bill himself. So, on the accommodated assumption that Bill's donkey ownership

[38]The circumstance that the descriptive content "rabbit" of the definite description *the rabbit* matches the constraint "rabbbit(y)" on the discourse referent y of (152) and that it doesn't match the descriptive constraint of the other discourse referents in the universe of (152), is enough for the interpreter to "zero-in" on y as the antecedent for u. It might be thought that such an interpretation carries with it the accommodation that the individuals represented by the other discourse referentes are not rabbits. But this isn't always so. For instance, consider *A man went to see a doctor. The doctor asked the man what was wrong with him.* Interpreting *the man* as anaphoric to *a man* doesn't carry the implication that the doctor is not a man. Nor does interpreting *the doctor* as anaphoric to *a doctor* carry the implication that the man who came to see him wasn't a doctor too.

involves a sole donkey, the pronoun should be just as effective in this case in selecting the intended antecedent as the description. But apparently it isn't.

This example points to a conclusion to which many other case studies point as well: An antecedent for a pronoun must have been introduced explicitly into the discourse beforehand; definite descriptions are happy to pick up entities whose existence is implied by the context, even if no explicit introduction has previously taken place.

This difference between pronouns and definite descriptions indicates that a uniform treatment of the two will go only so far. A theory which treats both pronouns and descriptions as anaphoric NPs will need an additional component which articulates the "anaphora resolution principles" according to which the antcedents for these two different NP types are determined. The analysis of definite descriptions we have exemplified in our treatment of (142.a) and (142.b) could be seen as a step in this direction. According to that analysis the anaphoric dimension of definite descriptions is located entirely in the determination of the contextual predicate C. However, if this is the way in which we want to make fully explicit precisely how descriptions differ from pronouns, then we will have to say much more about the rules according to which C may be resolved. This is arguably the central task for a theory of Bridging. It is a task on which some progress has been made in recent years, but which surely isn't yet completely solved. (See [Heim, 1982], [Bos *et al.*, 1995], [Clark, 1997], [Asher and Lascarides, 1998])

Summarising: It remains a question for further research exactly to what extent the analyses of definite descriptions and pronouns can be unified. We have seen that treating both as triggers of anaphoric presuppositions shifts the burden to articulating the different principles which govern the resolution of these presuppositions. Analysing definite descriptions, in the spirit of the logical tradition, as triggers of contextualised existence-and-uniqueness presuppositions brings out the difference between them and pronouns more clearly in principle, but work remains to be done as regards the contextual resolution of the anaphoric predicate variable C.

We have spent what may seem a disproportionate amount of space in this section on the analysis of pronouns and definite descriptions, and especially on the question how similar or dissimilar their analyses ought to be. Our justification for this is twofold. First, the analysis of pronouns and definite descriptions is a matter that has been of central importance in the philosophy of logic and language for well over a century. Second, the light in which these two NP types appear from the Dynamic perspective is radically different from the traditional picture, according to which anaphoric pronouns are variables and definite descriptions some species of referential term. (This is a view which, if we are not mistaken, is still prevalent among many philosophers and philosophical logicians today.) According to this view the two kinds of expressions are very different indeed. The Dynamic approach, however, makes it possible to see that, certain remaining discrepancies notwithstanding, the conceptual differences are far smaller than the traditional picture implies. All in all Dynamic Semantics (at least in the form in which it is being used here) projects a very different image of the way in which reference contributes to the expression of propositions, and to the range of posibilities of expressions which serve a referential role within the setting of dynamic interpretation.

Local vs global Justifcation

The examples we consider in this section are variants of those considered in Section 4.2. They have been chosen to show the difference between global and local justification of presuppositions. We will focus on just two logical sentence types, a conditional sentence (158.a) and a universal quantification (158.b).

(158) a. (It is a peculiar fact, but) If a friend of mine has both a rabbit and a guinea pig, the rabbit is white.[39]

b. Every friend of mine who has a rabbit overfeeds it.

The point of these examples is that the presuppositions associated with the definite description of (158.a) and the pronoun of (158.b) do not seem to "project". They don't, because they can be resolved in local contexts that are furnished by some other part of the sentence (the antecedent of the conditional or the restrictor of the quantification) than the part in which they are triggered (the conditional's consequent or the quantifier's nuclear scope). We note in passing that the *it* of (158.b) is a typical donkey pronoun.

The analysis of (158.a) resembles in most respects the one we gave in the last section of (142.a). The difference is that the presupposition is now adjoined to some embedded part of the preliminary representation (the consequent of the conditional), since it is this part that represents the sentence constituent which contains its trigger.

Adopting the same treatment of the definite description *the rabbit* which we used in our treatment of (142.a) we get as preliminary sentence representation the structure in (159), where $K_{(147)}$ stands for the preliminary DRS in (147).

(159)
$$\boxed{\begin{array}{c} \text{w} \\ \text{Walter(w)} \\ \boxed{\begin{array}{c} \text{y} \quad \text{z} \\ \text{rabbit(y)} \quad ; \quad \text{guinea pig(y)} \\ \text{have(w,y)} \quad ; \quad \text{have(w,z)} \end{array}} \Rightarrow \mathrm{K}_{(147)} \end{array}}$$

[39]The initial part of (158.a) in parentheses is intended to make the sentences a little less implausible, but the discussion will not take it into acount.

Note that the discourse referent w introduced by the NP *Walter* has been placed in the universe of the main DRS, not in that of the sub-DRS representing the antecedent of the conditional. In the original top-down construction algorithm for DRT (See Section 2) this was the effect of the processing rule for proper names. But in the present setting, where we have already committed ourselves to presuppositional accounts of pronouns and descriptions, it is natural to adopt an account also for other types of definite NPs, such as proper names and demonstratives. For each of these NP types the account must include an articulation of the resolution rules for the presuppositions generated by the NP in question. In the first part of this subsection we noted in relation to descriptions that determining what these resolution rules is anything but a trivial task, which a large spectrum of earlier investigations into the semantic and pragmatic behaviour of the given NP type requires taking account of. For other types of definite NPs the situation is no different. This is true in particular for proper names, whose referential properties have been the subject of countless publications within the philosophy of language, and, more recently, also within linguistics.[40] The fact that the referent of a proper name occurring in a discourse always finds a representation within the global context should come out as a consequence of the resolution rules for the presuppositions triggered by proper names.

We forgo an explicitly presuppositional treatment of proper names[41] here, and simply assume that the ultimate result of their interpretation is always representation at the level of the main universe of the sentence or discourse representation. There is one aspect to the use of proper names, however, which should be mentioned here. This is the ease with which they (or, in the present terminology, their presuppositions) are "accommodated". We have made passing mention of accommodation in the last section, and will have cause to do so a few more times until Section 4.4, in which accommodation is the topic. What matters right now are two points: (i) What sorts of information may be accommodated for the sake of justifying a given presupposition, and under what conditions, varies from one type of presuppositon to the next (cf. the discussion of bridging as a form of accommodation that is permissible in the case of definite descriptions but not of pronouns); and (ii) Accommodation is particularly unproblematic in the case of proper names: In cases where the interpreter of an occurrence of a proper name is unfamiliar with that name, or believes himself to be unfamiliar with the name's referent, he will normally assume that the speaker who is using the name knows who she is talking about. Accordingly he will accommodate

[40] Much of the philosophical literature of the past thirty years was in reaction to Kripke's Naming and Necessity [Kripke, 1972]. For an analysis of proper names that is directly relevant to the presuppositional account that is at issue here, see [Geurts, 1997].

[41] Likewise for the various types of demonstratives. Demonstratives, however, won't occur in any of the examples we discuss in this survey.

his interpretation context so that it contains a representation of the intended referent of the name, to which the discourse referent introduced by the given name occurrence can then be linked. We will assume henceforth that a name which does not as yet have a representation in the context will automatically lead to an accommodation of this kind.

Justification of the presupposition of (159) whose representation is adjoined to the representation of the consequent of the conditional can make use of the local context information that is provided by the antecedent — just as justification of the presuppositions of (147) could make use of the discourse context (143) provided by the preceding sentence. It is impotant to note, however, that the local context information is not restricted to the contents of the relevant sub-DRS itself, but includes also all information that belongs to other DRSs which are accessible from the local context. Thus, in the case at hand the discourse referent w and the condition "Walter(w)", which are not part of the sub-DRS representing the antecedent of the conditional, but of the main DRS, which is accessible from this sub-DRS, are part of the local context information too. It follows from this that if one context is more local than another (i.e. the second is accessible from the first, but not conversely), the first will always contain at least as much information as the second, and usually more.

This means among other things that resolution of C in the local context provided by the antecedent of the conditional in (159) can take the same form as it did in our treatment of (142.a), viz. that of identifying C with the predicate "$\in\{w,y,z\}$". After this identification, the effect of which results in replacing the DRS-component $K_{(147)}$ in (159) with $K_{(148)}$, satisfaction of the existence-and-uniqueness presupposition of the definite description follows as before and we end up with the presupposition-free sentence representation in (160).

$$(160)\quad \boxed{\begin{array}{c} \text{w} \\ \text{Walter(w)} \\ \boxed{\begin{array}{c} \text{y}\quad \text{z} \\ \text{rabbit(y)}\ \ ;\ \ \text{guinea pig(y)} \\ \text{have(w,y)}\ \ ;\ \ \text{have(w,z)} \end{array}} \;\Rightarrow\; \boxed{\begin{array}{c} \text{v} \\ \text{rabbit(v)} \\ \text{v}\in\{\text{w,y,z}\} \end{array}} \end{array}}$$

Inasmuch as the presuppositions of (159) have been justified on the strength of information provided by the antecedent of the conditional alone, we have accounted for why these projections do not project to become presuppositions of the conditional sentence as a whole. However, we should recall in this connection the observation made in the last section regarding our treatment of (142.a): The justifcation of the existence-and-uniqueness presupposition rests strictly speaking not only on the information explicitly provided, but also on further assumptions, such as that Walter is a human being and that human beings are not rabbits. Because of this it isn't strictly speaking true that the local justification of the presuppositions of (158.a) described here makes them disappear without any trace whatever. In a context in which these assumptions could not be made, (158.a) would impress the interpreter as infelicitous. However, for the sentence in question such contexts are unlikely and in general, there will be a presumption, shared by speaker and interpreter, that the context is not like this. Because of this general default assumption a sentence like (158.a) will appear to us as free of presuppositional constraints on context altogether.

Much the same treatment as the one we just presented for (158.a) is possible for (158.b). The presupposition triggered by *it* is now adjoined to the right hand side DRS of the duplex condition introduced by the quantifier *every friend of mine*. (161) gives the preliminary representation of the sentence and (162) the final sentence representation.

(161)
$$\begin{array}{|c|}\hline \text{m} \\ \text{speaker(m)} \\ \begin{array}{|c|}\hline \text{x y} \\ \text{friend-of(x,m)} \\ \text{rabbit(y)} \\ \text{have(x,y)} \\ \hline \end{array} \left\langle \forall \atop \text{x} \right\rangle \begin{array}{|c|}\hline \left\langle \left\{ \begin{array}{|c|}\hline \underline{\text{u}} \\ \text{non-pers(u)} \\ \hline \end{array} \right\}, \begin{array}{|c|}\hline \\ \text{overfeed(x,u)} \\ \hline \end{array} \right\rangle \\ \hline \end{array} \\ \hline \end{array}$$

(162)
$$\begin{array}{|c|}\hline \text{m} \\ \text{speaker(m)} \\ \begin{array}{|c|}\hline \text{x y} \\ \text{friend-of(x,m)} \\ \text{rabbit(y)} \\ \text{have(x,y)} \\ \hline \end{array} \left\langle \forall \atop \text{x} \right\rangle \begin{array}{|c|}\hline \text{u} \\ \text{u = y} \\ \text{overfeed(x,y)} \\ \hline \end{array} \\ \hline \end{array}$$

Examples involving presupposition triggers distinct from definite noun phrases

In this section we look at examples which involve presupposition triggers that are not NPs. The first of these is a factive verb, *be surprised that*, and the second the adverb *again*. We begin with sentence (163).

(163) Bill is surprised that he is late.

Factive verbs presuppose the truth of their clausal complements. Within the setting of the present account this means that in the preliminary representation the representation of the complement sentence must occur more than once — first, as argument of the attitudinal predicate expressed by the verb, and, second, as factive presupposition. This need for representation duplication is extremely common. It arises in all cases of factive presupposition and with many other presupposition triggers as well. Moreover, the duplication problem doesn't arise just in the context of presupposition. It is equally important in connection with ellipsis, and in that connection it has received a good deal of attention.[42] Here we will only be concerned with aspects of the problem that are specific to its manifestations in the context of presupposition.[43] Duplication poses a major problem: the duplicate representations must identify the same content. Strictly speaking we are dealing wwith just *one* interpretation of the sentence material for which duplicate representations seem required; the duplicate representations must all capture the content captured by *that* interpretation.

One way in which this duplication problem manifests itself has to do with the interpretation of pronouns occurring within a sentence or sentence part of which duplicate representations are needed. This form of the problem is illustrated by (163), where the anaphoric presuppositon trigger *he* occurs as part of the representation serving as second argument of the complement of *be surprised* and also as part of the factive presupposition.

To focus more clearly on the problem, let us consider a preliminary representation for (163) in which the *he*-presupposition is represented twice.

(164)
$$\boxed{\begin{array}{c} \mathrm{b} \\ \mathrm{Bill(b)} \\ \left\langle \{\mathrm{K}\}\ ,\ \boxed{\text{be-surpr.(b,}\left\langle \left\{ \boxed{\begin{array}{c} \underline{\mathrm{u}} \\ \mathrm{male(u)} \\ \mathrm{person(u)} \end{array}} \right\} ,\ \boxed{\text{“be-late”(u)}} \right\rangle)} \right\rangle \end{array}} \ ,$$

$$\text{where K is the DRS } \left\langle \left\{ \boxed{\begin{array}{c} \underline{\mathrm{u}'} \\ \mathrm{male(u')} \\ \mathrm{person(u')} \end{array}} \right\} ,\ \boxed{\text{“be-late”(u}'\text{)}} \right\rangle$$

The first thing to observe about (164) is that there will be no way to justify the factive presupposition unless we assume that (163) is used in a context

[42]See e.g. [Schiehlen, 1999; Schiehlen, July 2002] and the references cited there.

[43]The requirement that multiple representaions of the same sentence constituent express the same content is another reason for preferring a bottom-up contruction algorithm to one which works top-down. When we work our way bottom up, then normally we will have constructed a represention of the part which requires mutliple representation at the point when duplicates of the representation must be introduced into the representation that is being constructed. Disambiguation decisions that sometimes have to be made in the course of representation construction — we assume that syntactic trees may contain ambiguities which are resolved only when they are converted into semantic representations — will already have been made in this case. When we proceed top-down, copying will usually be needed at a point where the relevant part of the syntactic tree has not yet been converted. Special provisions have to be made to make sure that afterwards the same disambiguation decisions will be made in each of the copies.

that entails it. However, factive presuppositions are, like those connected with proper names, easily accommodated, and we will assume that this is what happens eventually in this case. But before the factive presupposition can be accommodated it ought to be clear what it is. In the present case this requires that its satellite presupposition, involving the anaphoric discourse referent u′, has already been resolved. So it is to the resolution of this presupposition, and of its alter ego involving the discourse referent u, that we must turn first.

When (163) is considered out of context — or against the background of an empty discourse context, as when it is the very first utterance of a conversation — the difficulty that multiple copies of anaphoric presuppositions can cause doesn't become visible yet. For in such a situation only Bill is available as anaphoric antecedent for both u and u′. Resolving the two pronoun presuppositions accordingly leads to the representation in (165).

(165) $$\begin{array}{|c|}\hline \text{b u}'\text{ u} \\ \text{Bill(b)} \quad ; \quad \text{u}' = \text{b} \\ \text{“be-late”(u}'\text{)} \\ \text{be-surpr.(b,} \begin{array}{|c|}\hline \text{u} \\ \text{u} = \text{b} \\ \text{“be-late”(u)} \\ \hline \end{array} \text{)} \\ \hline \end{array}$$

In (165) the representation of the second argument of *be-surprised* and that of the factive presuppostion do express the same proposition, viz the *de re* proposition which asserts of Bill that he is late. This is as it should be. But we cannot count on being so lucky always. Things may go wrong when a sentence like (163) is used in a non-empty context like that which is provided by the first sentence of (166).

(166) John was late and thats what he told Bill. Bill isn't surprised that he was late.

The context established by the first sentence of (166) contains the information that John was late and that John has told Bill that this is so. Among other things it introduces (representations for) John and Bill. So the pronoun *he* in the second sentence is now ambiguous between an interpretation in which it refers to Bill and one in which it refers to John. Once again we assume that the preliminary representation (now for the second sentence of (166)) is as in (164). Suppose we were to resolve u′ to John and u to Bill. This would give us on the one hand satisfaction of the factive presupposition in the context due to the first sentence (since the first sentence sasserts that John was late). On the other hand the non-presuppositional part of the sentence representation now says that Bill isn't surprised that Bill was. It is plain, however, that the factive presupposition of *that* claim isn't justified by the first sentence of (166). So the presupposition justification we obtain in this way is spurious. Evidently it won't do to resolve u and u′ to different antecedents.

There are various ways in which the requirement of a common resolution for the discourse referents u and u′ of the duplicate presupposition representations can be secured. One possibility is to insist that different copies of the same anaphoric presupposition all use the same distinguished discourse referent, and then to insist that different such occurrences (i.e. underlined occurrences) of the same discourse referent in the same representation just get the same antecedent. In the case of (164) this would mean that instead of the two discourse referents u and u′ we would have just one — u, say – occurring as distinguished discourse referent in both.)

These remarks are no more than suggestive. They can be turned into someh]thing more precise only within the context of an explicite construction algorithm for preliminary representations. This is of course a very important part of the theory of presupposition that is presented in this section. And not only that, it is a crucial part of the version of DRT presented here, given that presuppostion (with anaphors as special case of it) is the central phenomenon that DRT should be able to account for. Unfortunately a systematic treatment of DRS-construction — the construction of preliminary sentence representations and the integration of these into discourse representations — would transcend the already strained bounderies of the present chapter. We will say a few words about the two phases of DRS construction in Section 4.4 and 4.5, respectively. For a more detailed treatment we point the reader to the forthcoming [Hans and Reyle,].

We conclude this discussion with one last remark. on factive presuppositions here. As a rule factive predicates serve the purpose of attributing propositional attitudes. Such predicates do not only give rise to factive presuppositions, but also to presuppostions to the effect that the content of the attributed attitude is something that the attributee himself believes (or even takes for granted). At this point we are lacking the means to represent such "doxastic" presuppositions adequately. Once we will have procured these means in Section 5, we will return to the presuppositions of factive verbs.

Our next example involves the presupposition trigger *again*. *again* triggers a presupposition to the effect that the event or state described by the sentence constituent to which it is adjoined was preceded by an event or state satisfying the same description.[44] To account for such presupposi-

[44]This is an oversimplification. As has been noted by several authors [Fabricius-Hansen, 1980; Fabricius-Hansen, 1983; Stechow, 1996], *again* is ambiguous between a *repetitive* and a *restitutive* interpretation. The difference is most clearly seen with certain telic verbs, for instance, the verb *cure*. In *The tourist came down with typhoid, but the local doctor cured him again.* The word *again* can either be interpreted as presupposing that

tions we must be able to represent the temporal aspects of natural language meaning. So we now need as our basic DRT-language, which provides the building blocks from which preliminary representations are built, one in which events, states and their temporal properties and relations are made explicit. We take as underlying language the DRS language of Section 3.5. (See also [Kamp and Reyle, 1993].)

First consider the sentence (167), with the preliminary representation in (168).[45]

(167) John made a mistake again.

(168)
$$\begin{array}{|c|}\hline \mathrm{j} \\ \mathrm{John(j)} \\ \left\langle \left\{ \begin{array}{|c|}\hline \mathrm{t'\ e'\ y'} \\ \mathrm{mistake(y')} \\ \mathrm{e'{:}make(j,y')} \\ \mathrm{e'} \subseteq \mathrm{t'} \\ \hline \end{array} \right\}, \begin{array}{|c|}\hline \mathrm{t\ e\ y} \\ \mathrm{mistake(y)} \\ \mathrm{e{:}make(j,y)} \\ \mathrm{e} \subseteq \mathrm{t} \\ \mathrm{t'} \prec \mathrm{t} \\ \hline \end{array} \right\rangle \\ \hline \end{array}$$

The aspect of (168) that requires more extensive discussion is the placement of the condition "$t' \prec t$". This condition expresses that the presupposed eventuality precedes the asserted one and this is clearly needed. But where should we put it? In (168) it has been inserted into the non-presuppositional part of the representation, to which the *again*-presupposition has been adjoined. The reason for doing this is that a presupposition has a certain "logical priority" over the sentence or sentence part to which it is adjoined: The semantic contribution that is made by this sentence or sentence part will be determined only when the presuppostion has been resolved. So the presuppotion should not be "referentially dependent" on the representation of its adjunction site in the sense of containing a free discourse referent which is bound only in the representation of the adjunction site. But adding the condition "$t' \prec t$" to the representation of the presupposition would create just such a dependency. Hence its appearance as part of the representation of the adjunction site.

---

the there was an earlier event of the doctor curing the patient (the repetititve reading), or as presupposing that before the time when the tourist came down with typhoid he was in a state of being healthy (or at least typhoid- free) and that this state of affairs is "restituted" by the curing event whose occurrence is asserted by the sentence (the restitutive reading). Here we will only consider repetitive readings.

[45]The representation given here of the VP *make a mistake* is not really satisfactory. First, it isn't right to analyse make as a relation between the subject Bill and some independently existing object, the mistake. *make* functions as "verb of creation" here and the mistake is what results from the event it describes. Second, *make* acts as a light verb. The relation it contributes cannot be determined from the verb by itself, but only in combination with the head noun *mistake* of its direct object.

There is, however, another intuition which seems to speak against adding "$t' \prec t$" to the adjunction site representation. *Again*-sentences like (167) seem to be making a certain claim, expressed by the sentence without *again*. *Again* seems to "tag on" a certain presupposition to this claim, to the effect that one or more eventualities of the kind described by the claim occurred before the one whose occurrence is claimed. From this perspective the presuppostion is that one or more eventualities of the kind described occurred *before the occurrence time of the asserted eventuality*: The latter eventuality ev is said to have occurred at some time t and the presupposition is that there were similar eventualities before t. This perspective is especially compelling in connection with certain negated *again*-sentences. For instance the second sentence of

(169) Mary came on Tuesday. But she didn't come again on Wednesday.

is naturally glossed as (i) making the claim that there was no coming on Wedensday, and (ii) that there was a coming of Mary at some time before Wednesday (a presupposition which is justified by the first sentence of (169)).

But for non-negated sentences such as (167) the perspective is plausible as well. It is easy to imagine a context in which a certain past time is already in focus and in which the ..... of (167) is understood to locate the event it describes. In fact, as we saw in Section 3.5, tenses are often anaphoric in several ways; now that we have reinterpreted anaphora as a species of presuppositions we can make this anaphoric dimension more explicit by representing the location time t of the asserted event as involving an anaphoric presupposition of its own. In a preliminary representation in which t is treated in this way, the *again*-presupposition can now be made dependent on this first presupposition, without it thereby becoming dependent on the non-presuppositional part of the representation. Such a preliminary representation is given in (170).

(170)
$$\begin{array}{|c|}\hline \mathrm{j} \\ \mathrm{John(j)} \\ \left\langle \left\{ \begin{array}{|c|}\hline \underline{\mathrm{t}} \\ \mathrm{t} \prec \mathrm{n} \\ \hline \end{array} \right\}, \left\langle \left\{ \begin{array}{|c|}\hline \mathrm{t}'\ \mathrm{e}'\ \mathrm{y}' \\ \mathrm{mistake(y')} \\ \mathrm{e}'\mathrm{:make(j,y')} \\ \mathrm{e}' \subseteq \mathrm{t}' \\ \mathrm{t}' \prec \mathrm{t} \\ \hline \end{array} \right\}, \begin{array}{|c|}\hline \mathrm{e\ y} \\ \mathrm{mistake(y)} \\ \mathrm{e:make(j,y)} \\ \mathrm{e} \subseteq \mathrm{t} \\ \hline \end{array} \right\rangle \right\rangle \\ \hline \end{array}$$

An interesting feature of *again*, which it shares with anaphoric words such as *else*, *other* and relational adjectives such as *similar*, *related* etc., is shown by the following pair of discourses.

(171) a. I will help Bill tomorrow. But I won't help him again.

b. I will help Bill tomorrow. But I will not help him.

c. I will help Bill tomorrow. But I will never help him again.

d. I will help Bill tomorrow. But I will never help him.

(171.b) and (171.d) are bizarre and can only be interpreted as straight contradictions: the first sentence announces that a certain event will take place tomorrow and the second asserts that no such event will take place (ever) in the future. (171.a) and (171.c) are perfectly coherent. They convey that a certain event will occur tomorrow and that no such event will ever happen *after this one.* Evindently it is the presence of *again* in (171.a,c) and its absence from (171.b,d) which is responsible for the difference.

To account for this contrast we must recall what was said about frequency adverbs and negation in 3.5. The interpretation of both negation and adverbs like *never* involves a frame adverbial, we noted there. This frame adverbial plays the same role in such sentence as does the location time in simple sentences like (167) or the first sentence of (171.a,b); in particular, in sentence with *not* or *never* it is now the frame interval that is involved in the anaphoric location time presupposition displayed in (170). Thus we get for the second sentence of (171.a) which contains *not* the representation in (ref29"), whereas (ref29"') is the representation of the corresponding sentence without *again*.

(172)
$$\begin{array}{|c|}\hline \text{j b} \\ \text{the\_speaker(j)} \;\;;\;\; \text{Bill (b)} \\ \left\langle \left\{ \begin{array}{|c|}\hline \underline{\text{t}} \\ \text{n} \prec \text{t} \\ \hline \end{array} \right\}, \left\langle \left\{ \begin{array}{|c|}\hline \text{t}' \ \text{e}' \\ \text{e}'\text{:help(j,b)} \\ \text{e}' \subseteq \text{t}' \\ \text{t}' \prec \text{t} \\ \hline \end{array} \right\}, \begin{array}{|c|}\hline \text{s} \\ \text{dur(s)} = \text{t} \\ \text{s:}\neg \begin{array}{|c|}\hline \text{e} \\ \text{e} \subseteq \text{t} \\ \text{e:help(j,b)} \\ \hline \end{array} \\ \hline \end{array} \right\rangle \right\rangle \\ \hline \end{array}$$

(173)
$$\begin{array}{|c|}\hline \text{j b} \\ \text{the\_speaker(j)} \;\;;\;\; \text{Bill (b)} \\ \left\langle \left\{ \begin{array}{|c|}\hline \underline{\text{t}} \\ \text{n} \prec \text{t} \\ \hline \end{array} \right\}, \begin{array}{|c|}\hline \text{s} \\ \text{dur(s)} = \text{t} \\ \text{s:}\neg \begin{array}{|c|}\hline \text{e} \\ \text{e} \subseteq \text{t} \\ \text{e:help(j,b)} \\ \hline \end{array} \\ \hline \end{array} \right\rangle \\ \hline \end{array}$$

The contrast between (172) and (173) shows that the presence of the *again*-presupposition in (171) creates the possibility of resolving t' to the time t_0 of the event of the speaker helping Bill tomorrow which is asserted in the first sentence and resolving t to the period following t′ (=t_0). In (173), where the *again*-presupposition is absent, such a resolution of t is apparently not possible (even though the context is exactly the same). We will return to this case in Section 4.5.

There is one further feature of (172) which requires comment here. This is the interaction between *again*-presuppositions and negation.[46] Even though syntactically *again* is presumably within the scope of negation in (171.a) the presupposition it triggers has been adjoined outside the negation operator in (172). What is the justification for this?

The answer is connected with an aspect of the theory of presupposition that was prominent from the first beginnings. This is the relation between presupposition and negation. (The interaction between negation and *again*-presupposition in (171) is just one instance of this; but the relation equally concerns all type of presuppositions.) In the early days, when presupposition was primarily the concern of logicians it was seen as one of the central features of presuppositions that they equally affect a sentence S and its negation. The presuppositions of ¬S are the same as the presuppositions of S. If the presuppositions are not satisfied, then neither S nor ¬S is (i.e. have a truth value, or express a proposition); if the presuppositions are satisfied, then both S and ¬S are "proper" (and, of course, of opposite truth value). When presupposition theory became a part of linguistics later on, the question whether an implication is preserved under negation became one of the major criteria for deciding whether the implication is a case of presupposition.

Whithin the setting of present account of presupposition, the fact that presuppositions of S are also presuppositions of ¬S has a simple explanation. Negation is a 1-place operator. (In this regard it differs both from

[46]Note that the scope relation between the negation and *again* in the second sentence of (174.a) is also evident semantically. When *again* has scope over a negation, then the negation will figure in the *again*-presupposition. Compare for instance (174.a) – (174.d)

(174) a. Bill was on time yesterday. But he hasn't been on time again today.

b. Bill wasn't on time yesterday. And he hasn't been on time again today.

c. Bill was on time yesterday. *But again he hasn't been on time today.

d. Bill wasn't on time yesterday. And again he hasn't been on time today.

These examples make two points. First, the oddity of (174.c) shows that when *again* unequivocally has scope over the negation, then the negation is part of the presupposition. The *again*-presupposition of the second sentence of (174.c) and (174.d) says that there was an earlier occasion when Bill wasn't on time. This is precisely what the first sentence of (174.c) does not allow us to justify. (The first sentence of (174.d), on the other hand, does allow this, so (174.d) is unobjectionable.) That both (174.a) and (174.b) are fine shows that there the second sentence is ambigous between a reading in which *again* has wide scope over the negation and one in which it does not. In our representation of (174.a) we have been guided by the semantic intuition that here the narrow scope reading for *again* seems the intended one.

"sentential connectives", like *and* and *if*, and from natural language quantifiers, like *all*, *most*, *many*, etc., all of which are 2-place (the quantifiers always involve restrictor as well as nuclear scope). As a consequence there is within the scope of the negation no new information that could serve as local context for the justification of a presupposition triggered by the sentence or sentence part on which the negation operates (in the way in which the antecedent of a conditional can serve as context for the justification of a presupposition triggered by the consequent, etc.) So justification of the presupposition will be possible, if at all, only in a context which includes the negation within its scope.

In view of this it is immaterial for the final outcome of the interpretation process whether a presupposition triggered within a negated sentence or sentence part is adjoined under or outside the scope of the negation operator. For instance, we could have represented the second sentence of (171.a) also as in (175), with the representation of the presupposition inside the scope of $\neg$.

$$(175)\quad \boxed{\begin{array}{c} \mathrm{j\ b} \\ \mathrm{the\_speaker(j)}\ \ ;\ \ \mathrm{Bill\ (b)} \\ \left\langle \left\{ \boxed{\begin{array}{c} \underline{\mathrm{t}} \\ \mathrm{n} \prec \mathrm{t} \end{array}} \right\}, \boxed{\begin{array}{c} \mathrm{s} \\ \mathrm{dur(s)} = \mathrm{t} \\ \mathrm{s}{:}\neg \boxed{ \left\langle \left\{ \boxed{\begin{array}{c} \mathrm{t}'\ \mathrm{e}' \\ \mathrm{e}'{:}\mathrm{help(j,b)} \\ \mathrm{e}' \subseteq \mathrm{t}' \\ \mathrm{t}' \prec \mathrm{t} \end{array}} \right\}, \boxed{\begin{array}{c} \mathrm{e} \\ \mathrm{e} \subseteq \mathrm{t} \\ \mathrm{e}{:}\mathrm{help(j,b)} \end{array}} \right\rangle } \end{array}} \right\rangle \end{array}}$$

In the context provided by the first sentence of (171) justification of the *again*-presupposition of (175) clearly comes to the same thing as justification of the *again*-presupposition of (172): In both cases justification is in the global context specified by the first sentence, so exactly where the presupposition is adjoined in the preliminary representation of the new sentence is, as far as justification is concerned, of no consequence. The same is true in cases where the presupposition is justified sentence-internally. (Consider, e.g., the conditional with the first sentence of (171.a) as antecedent and the second sentence as consequent.) Here too it doesn't matter whether the justified presupposition is represented under or outside the negation operator.

Is there any way of choosing between the preliminary representations (172) and (175)? Arguably it is (175) which results directly from the application of the general construction rules to the input structure (i.e. the syntactic structure of the second sentence of (171.a)). (172) can then be seen as a variant of (175) which can be obtained by "lifting" the presupposition "beyond" the negation, relying on the principle that justification of the presupposition is equivalent to justification in a context which has scope over the negation.

(175) is preferable to (172) also for another reason. Sometimes presupposition can be "cancelled under negation". This typically involves denying explicitly that the presupposition is true. A famous example which played its part in the debate between Russel and Strawson over the question whether there is such a thing as presupposition is given in (176.a); (176.b) illustrates the same phenomenon, but in connection with a presupposition triggered by *again*.

(176) a. The exhibition wasn't opened by the King of France.

b. I am not reading this paper again. I am reading it for the very first time.

Sentences of this kind are described as instances of "local satisfaction of the presupposition under negation", a term which suggest that what we see in these cases is on a par with non-projection of the kind exemplified by, for instance, ... We believe, however, that the two phenomena are very different in nature and that the right explanation of what we see in (176) is along the lines of Horn's theory of negation. According to this theory the function of negation is not restricted to denying the truth of the proposition expressed by the sentence material in its scope, but can be extended to cover other factors that can be responsible for failure to produce a felicitous true claim. Failure of presupposition is one of the many factors. There is an ongoing dispute precisely under what circumstances negation can be used to reflect factors other than the actual falsity of a correctly expressed proposition. (In order that a negation can be understood as denying truth of what it applies to because of presupposition failure, it seems neccessary that the presupposition must itself be denied explicitely, but probably the last word about what is involved in presupposition cancellation has not yet been said.) However, it should be clear, whatever the details, that presupposition cancellation is an entirely different phenomenon from local satisfaction. In the latter case the presupposition is "locally true"; in the former it is "globally false" and thus a fortiori "locally false" as well.

Local Justification in Conjunctions

In this subsection we briefly address one further instance of local presupposition justification, that of sentence compounds formed with "&-like" conjunctions such as *and*, *but*, *although*, *because* and others. It is one of the standard facts about anaphora that pronouns to the right of such a conjunction can be construed as anaphoric to an indefinite on the left, but not vice versa. This is so both when the compound is a complete sentence by itself and when it occurs as part of a larger sentence, e.g. as antecedent or consequent of a conditional. The same observations apply to presuppositions that are due to other triggers than pronouns. We restrict our attention here to coordinate structures with the conjunction *and*.

(177) a. If a friend of mine has a rabbit and he loves it, then he overfeeds it

b. *If he loves it and a friend of mine has a rabbit, then he overfeeds it

c. If someone is caught stealing and he is then caught stealing again, he is sent to goal.

d. *If someone is caught stealing again and he is then caught stealing, he is sent to goal.

From the point of presupposition justification conjunctions thus behave like conditionals and quantifications. It is natural to capture this similarity by representing conjunctions at the level of preliminary representations in the form of complex DRS conditions, composed of the representation for the first conjunct and the representation of the second. This will enable us to specify, by the same means that we used for duplex conditions and conditions formed with $\Rightarrow$, that the first conjunct serves as local context for the second conjunct.

To form such conjunctive DRS conditions we need an operator to represent the conjunction operator which English expresses by means of the word *and*. From the Dynamic Semantics literature it is clear what symbol we should use for this purpose. Dynamic Semantics, in the more restrictive sense of the term, makes use of the dynamic conjunction operator ";".[47] The semantics of this operator stipulates that a conjunction formed with its help is true in a given context K_0 iff (i) the first conjunct is true in K_0 and (ii) the second argument is true in the context obtained from updating K_0 with the information contributed by the first conjunct. This is in essence what we need. And in Section 4.3, where we present the model theory for preliminary DRSs, we will define the semantics of ";" along these lines.

Using ";" we get for (177.a) the preliminary representation in (178).

(178)
$$\boxed{\boxed{\boxed{\begin{array}{c} \text{x y} \\ \text{boy(x)} \\ \text{rabbit(y)} \\ \text{have(x,y)} \end{array}} \; ; \left\langle \mathcal{K}_1, \boxed{\text{love(z,u)}} \right\rangle} \;\Rightarrow\; \boxed{\left\langle \mathcal{K}_2, \boxed{\text{overfeed(w,v)}} \right\rangle}}$$

[47]See for instance [Groenendijk and Stokhof, 1991].

$$\mathcal{K}_1 = \left\{ \boxed{\begin{array}{c} \underline{\text{z}} \\ \text{male(z)} \\ \text{pers(z)} \end{array}}, \boxed{\begin{array}{c} \underline{\text{u}} \\ \text{non-pers(u))} \end{array}} \right\}$$

$$\mathcal{K}_2 = \left\{ \boxed{\begin{array}{c} \underline{\text{w}} \\ \text{male(w)} \\ \text{pers(w)} \end{array}}, \boxed{\begin{array}{c} \underline{\text{v}} \\ \text{non-pers(v))} \end{array}} \right\}$$

We assume that the left sub-DRS of the ;-condition in (178) can serve as local context for the justification of the presuppositions adjoined to its right sub-DRS. This allows for resolution of the anaphoric discourse referents z and u via identification with x and y. The result of this can be represented as in (179).

(179)
$$\boxed{\boxed{\boxed{\begin{array}{c} \text{x y} \\ \text{boy(x)} \\ \text{rabbit(y)} \\ \text{have(x,y)} \end{array}} \; ; \; \boxed{\begin{array}{c} \text{z u} \\ \text{z = x ; u = y} \\ \text{love(z,u)} \end{array}}} \Rightarrow \boxed{\left\langle \mathcal{K}_2, \boxed{\begin{array}{c} \\ \text{overfeed(w,v)} \end{array}} \right\rangle}}$$

$$\mathcal{K}_2 = \left\{ \boxed{\begin{array}{c} \underline{\text{w}} \\ \text{male(w)} \\ \text{pers(w)} \end{array}}, \boxed{\begin{array}{c} \underline{\text{v}} \\ \text{non-pers(v))} \end{array}} \right\}$$

What about the justification of the remaining presuppositions, in the consequent of the conditional? Here we face an issue which we have not yet encountered. By representing the antecedent of the conditional (177.a) in (178) as a complex condition inside the left DRS of the ⇒-condition we have created a configuration which seems to render the information inside the DRS components of the ;-condition inaccessible to that in the consequent of the ⇒-condition. Clearly this is not what we want. We could extend the definition of accessibility in such a way that the information in the conjuncts of ; in (179) does become accessible to the consequent. But the same effect can also be achieved in a slightly different way, viz. via the priciple that presupposition-free ;-conjunctions can be merged with the DRS to whose condition set the conjunction belongs. The principle is stated in (180).

(180) (Lifting of presupposition-free ;-conditions)

Suppose that a preliminary representation K has a component K′, that K′ contains a condition of the form "K_1;K_2" and that this condition is a DRS condition — that is, both K_1 and K_2 are free of presuppositions. Then K' may be replaced in K by the merge of K', K_1 and K_2.

Applying (180) to (179) gives (181), in which the remaining presuppositions can now be resolved in the manner desired. The final representation is (182).

(181) $$\boxed{\begin{array}{|c|}\hline \text{x y z u}\\ \hline \text{fr.o.m.(x)}\\ \text{rabbit(y)}\\ \text{z = x}\\ \text{u = y}\\ \text{have(x,y)}\\ \text{love(z,u)}\\ \hline\end{array} \;\Rightarrow\; \boxed{\left\langle \left\{ \begin{array}{c} \begin{array}{|c|}\hline \underline{\text{w}}\\ \hline \text{male(w)}\\ \text{pers(w)}\\ \hline\end{array} \\ \begin{array}{|c|}\hline \underline{\text{v}}\\ \hline \text{non-pers(v)}\\ \hline\end{array} \end{array} \right\}, \begin{array}{|c|}\hline \\ \hline \text{overfeed(w,v)}\\ \hline\end{array}\right\rangle}}$$

(182) $$\boxed{\begin{array}{|c|}\hline \text{x y z u}\\ \hline \text{fr.o.m.(x)}\\ \text{rabbit(y)}\\ \text{z = x}\\ \text{u = y}\\ \text{have(x,y)}\\ \text{love(z,u)}\\ \hline\end{array} \;\Rightarrow\; \begin{array}{|c|}\hline \text{w v}\\ \hline \text{w = x}\\ \text{v = y}\\ \text{overfeed(w,v)}\\ \hline\end{array}}$$

Applications of principle (180) seem to introduce into the justification process a genuine element of sequentiality. For instance, in (179) we must first resolve the presuppositions involving z and u, then apply (180) and then resolve the remaining presuppositions. In this regard the processing order that is imposed on (179) differs from the cases discussed in the last section. There too, presupposition resolution often seemed to be a matter of preceeding in the right order. For instance, in discussing the *again*-sentence (167) we attributed a logical priority to the resolution of the presuppostion for the location time t over the presupposition triggered by *again*. But it is nevertheless possible to understand presuppostion resolution for a given preliminary representation as a single problem, that of the simultaneous solution of a set of (presuppositional) constraints: each constraint resolution must be chosen in such a way that all fit together. In Section 4.5 we will look at examples which bring out this aspect of simultaneous constraint solving very clearly.

Principle (180) appears to change this picture. Usually it requires that certain resolutions have already been carried out while other resolutions are possible only once the application has occurred. But of course this doesn't alter the spirit of simultaneous resolution. Even in the presence of (180) it remains true that the earlier resolutions should be carried out in such a way that the later ones remain possible. In fact, we can eliminate the element of sequentiality which the introduction of (180) introduces by redefining accessibility (as suggested earlier), and leaving (180) merely as a means for simplifying the notation of presupposition-free representations. But this is an alternative we won't pursue further here.[48]

We noted that *and* is only one of a number of words that form logical conjunctions in English. In some cases, like in that of *and* itself these conjunctions take the form of coordinations and in others (e.g. *because*) that of subordinations; but in all these cases left-to-right order matters to anaphoric an presupposition resolution, and therefore requires the use of ; in preliminary representation for the same reasons why it is needed in the representation of conjunctions with *and*. (There is a complication with subordinate conjunctions. When the subordinate clause precedes the main clause, resolution may "go in either direction", with the main clause serving as context for the justification of presuppositions arising within the subordinate clause as well as the other way round. These cases require a more complex analysis, which we do not go into here. (For the case of pronominal anaphora see [Chierchia, 1991].)) But it is not only coordinate and subordinate conjunctions which require the use of ;. The effect of linear order on presupposition resolution makes itself fell in many other . . . as well. One type of case involves relative clauses. An example is given in (183).

(183) A man who loves a woman who also loves him ought to be happy.

Here *him* can (and is naturally understood to) refer to the discourse referent introduced by *a man* and also is appropriate in that the proposition

[48]One reason for dwelling on the case of conjunction has to do with the history of DRT. The problem was known from the earliest days of DRT, at least in its application to anaphoric pronouns. In [Kamp and Reyle, 1993] it was discussed at considerable length (see Ch. 1.5). But the solution presented there, involving a baroque indexation system which takes away much of the initial appeal of DRSs as comparatively simple data structures, can be euphemistically described as "awkward". At that time the dynamic conjunction operator ; was already widely known and it was certainly known to the authors. The failure to make use of ; in [Kamp and Reyle, 1993] was based on a certain confusion: In DRSs the need for conjunction as a logical operator is rendered superfluous by the device of collecting DRS conditions into sets — the set Con_K consisting of the DRS conditions of the DRS K acts as the conjunction of those conditions. But sets are by definition unordered, so the left-right ordering between the conjuncts of a conjunction in natural language is lost as soon as those conjuncts, or their representations, are made into a set. The indexing system of [Kamp and Reyle, 1993] was designed to retain information about their order as long as it was needed, but the solution seems at hoc and is unappealing.

The two-stage architecture adopted here gives us the right way of having our cake and eating it. As indicated by our discussion of (174.a), we need the order within the preliminary representations, but once presupposition justification has taken place, it can be discarded. In the present formulation, it is principle (180) which does the discarding.

expressed by the clause to which it is adjoined — that of the woman y loving the man x — is parallel in the way required for the interpretation of *also* to the proposition that x loves y which "precedes" it. It is a well-known point of generative syntax that reversal of *him* and its antecedent distroys this interpretation and the same can be observed when we move *also* from the inner to the outer relative clause.

(184) a. He who loves a woman who also loves a man ought to be happy.

b. A man who also loves a woman who loves him ought to be happy.

(184.a) can only be interpreted as a statement about men who love a women who love some man or other, and presumably one that is different from themselves. And (184.b) implies that there is something else connected with the man (some other property, or something he does) besides loving a woman who also loves him; the parallel between x loves y and y loves x doesn't help in this case to justify the presence of *also*. If we want to capture these asymmetries with the help of ;, then the component DRSs of the restrictor of the generic quantification in (183) will have to be separated by ; at least to the extend shown in (185).[49]

(185)
$$\boxed{\boxed{\begin{array}{c} \mathrm{x} \\ \mathrm{man(x)}\ ;\ \boxed{\mathrm{K}} \end{array}}\ \ \begin{array}{c} \mathrm{x} \\ \forall \end{array}\ \ \boxed{\begin{array}{c} \\ \text{“ought to be happy”(x)} \end{array}}} \quad \text{with K} =$$

$$\boxed{\begin{array}{c} \mathrm{y} \\ \mathrm{woman(y)} \\ \mathrm{loves(x,y)} \end{array}}\ ;\ \left\langle \left\{ \boxed{\begin{array}{c} \underline{\mathrm{u}} \\ \mathrm{male(u)} \\ \mathrm{person(u)} \end{array}}\ ,\ \boxed{\begin{array}{c} \mathrm{v\ w} \\ \mathrm{v} \neq \mathrm{y} \\ \mathrm{w} \neq \mathrm{u} \\ \mathrm{loves(v,w)} \end{array}} \right\},\ \boxed{\mathrm{loves(y,u)}} \right\rangle$$

We leave it as an excercise to verify (i) that the two presuppositions of (185) can be resolved in the intended ways on the assumption that what comes to the left of ; can serve as context for what occurs to the right, and (ii) that the intended resolutions are blocked in similarly constructed preliminary DRSs for (184a,b).

For relative clause constructions the construction principle yields the ;-articulated representations we want. But the asymmetry problem is much more pervasive. (186) gives two further examples.

[49] The represenation of the *also*-presupposition is based on the assumption that the presupposition *also* generates is similar to the one expressed by *also*'s adjunction site. For this particular case we have assumed that such a proposition can take the form of a combination of the same relation, "love", with different arguments. A proper treatment of the presuppositions triggered by also requires an account of information structure which has not been included in this chapter; so the present treatment of the *also*-presupposition has to be taken at force value.

(186) a. A little boy loved his rabbit.

b. When a bishop meets another bishop, he blesses him.

In (186.a) switching the order of the two NPs distroys the possibility of interpreting *his* as anaphoric to *a little boy* (whether or not we keep the verb in the active, as in *His rabbit loved a little boy.*, or turn it into the passive, as in *His rabbit was loved by a little boy.*) Similarly switching subject and object in (186.b) — *When another bishop meets a bishop he blesses him.* — it becomes much harder (if perhaps not outright impossible) to interpret *another* as anaphorically related to the object NP *a bishop* (i.e. as understanding the latter NP as denoting an individual from which that of the NP *another bishop* is disjoint). Rather, the new sentence suggests that some bishop is already part of the discourse context.

These last two examples illustrate two points: (i) apparently the asymmetry conditions which limit the possibilites for presupposition resolution are not the same for presuppositions of different types. (The constraint on *his* seems to be stricter than that on *another.*); (ii) the conditions seem to depend on subtle questions of grammatical structure. For the special case of anaphoric pronouns this problem has been thoroughly investigated within the so-called Binding Theory of Government and Binding and subsequent frameworks of Chomskyan or Chomsky-inspired syntax.[50] For anaphoric presuppositions of other kinds and for non-anaphoric presuppositions (for the distinction between anaphoric and non-anaphoric presuppositions as we here understand it, see Section 4.2) the issue is, as far as we know still largely unexplored. As long as the empirical facts in this area remain in the dark there is no hope of stating the principles which guarantee that perliminary representations receive the correct ;-articulation. But, equally important, if the asymmetry constraints on presupposition resolution vary from one type of presupposition to the next, ;-articulation cannot be a sufficient guideline to presupposition resolution in any case. At least for some kinds of presuppositions the relevant structural constraints will have to be represented or defined in some other way. At the present point in time the range of questions that need to be answered in this domain is only gradually coming into sharper view. All we can do here is to point somewhat loosely in this general direction.

4.3 Syntax and Semantics of Preliminary Representations

The examples we have shown in Section 4.2 should have given some impressions of the form that preliminary sentence representations can take, but this impression is inevitably incomplete. The formal definition we present

[50]See, e.g., [Chomsky, 1981] and [Lasnik, 2003]. For a proposal how the Binding Theory of GB can be integrated into a method of constructing DRSs see [Berman and Hestvik, 1994].

in this section of the syntax of preliminary representations is quite liberal. It specifies a class of preliminary representations which includes not only the types exemplified in the last section, but much else besides. For all that is known to us at this point, the class may well be in excess of what it needs to be from the perspective of the semantics of a natural language. But this is an issue that it will be possible to settle only when the present proposal has been applied to a much larger range of cases than it has been so far.

We assume a DRS language L as given. DRS_L denotes the set of all DRSs of L. Exactly what L is like won't matter to the definitions which follow. We have by now come across quite a number of DRS-languages: the "basic" language used in Section 2 and formally defined in Sections 3.1 and 3.2, and the extensions that were introduced in Sections 3.3 — 3.5. Since the extensions proposed in 3.3, 3.4 and 3.5 are independent of each other, this already gives us a range of 7 possible extensions of the basic language. But in fact this is an underestimate since the extensions proposed in 3.3 and 3.5 were not uniquely determined; rahter, each can take one of a number of different forms, depending of the adopted set of quantifiers in the case of 3.3 or, in the case of 3.5, depending on the set of aspectual operators, and on the set of adverbial temporal quantifiers. The formal definitions which will follow should be independent of which of the possible DRS languages we choose for our underlying language L. To this end we assume that L is equipped with a certain set OP_L of "complex condition formers" — operators O^n — which build complex DRS conditions from n-place sequences of "argument" DRSs, while binding one or more discourse referents in the process (in the sense in which a quantifier like $\forall$ binds the discourse referent appearing in the central diamond of the corresponding duplex condition). We will assume for simplicity that no operator binds more than one discourse referent at a time (thus leaving the cases of polyadic quantification discussed at the end of Section 3.3 out of consideration). On this assumption it is possible to distinguish between the "variable binding" and the "non-variable binding" operators in terms of a binary feature: we mark each variable binding operator with a "+", writing O^n_+, while leaving the non-variable binding ones unmarked. We restrict the scope of the possible languages L in other ways as well, in that we ignore operators like PROG, which apply to intensional event abstracts over DRSs, as well as predicates like the attitudinal state predicate 'Att' of the next section whose second argument is an ADS (Attitude Description Set), an expression type of a complexity not yet encountered. Again these restrictions are not essential and are easily removed once the definitions are in place for the more restricted set of languages we will consider. We need one furhter peace of information about our operators, viz. the accessibility relation among their arguments. This is something that cannot be predicted in general terms — recall the difference between $\Rightarrow$ and $\vee$, with the first argument being accessible from the second (but not vice versa) in the case of $\Rightarrow$, but with no accessibility either way in the case of $\vee$. Since we take the accessibility relation within a preliminary DRS K to be a strict partial order (i.e. a relation which is transitive and asymmetric)[51] and since it will contain the accessibility relations between the components of a DRS condition $O^n(\mathrm{K}_1,...,\mathrm{K}_n)$ as a suborder in case K contains this condition, the accessibility relation among K_1, ..., K_n will have to be a strict partial order as well. We assume that the accessibility relation among the arguments of any operator $O^n_{(+)} \in OP_L$ is given as part of OP_L, and assume that this information is given in the form of a function Acc on OP_L which assigns to each $O^n_{(+)} \in OP_L$ a strict partial order on the set $\{1, ..., n\}$. for operators $O^n_{(+)}$ with more than two arguments we need to make two further assumptions relating to their local accessibility relation, Acc_{O^n}. First, in order that the definitions below are well behaved it is neccessary to assume that for each argument K_j of a condition $O^n(\mathrm{K}_1,...,\mathrm{K}_n)$ from which at least one other argument K_i of the condition is accessible, there exists a "nearest accessible" argument.

(187) Suppose that $\langle$i,j$\rangle \in Acc_{O^n}$, then there must be among the argument positions 1,...,n one position k which is a "minimal predecessor of j in Acc_{O^n}", i.e. $\langle$k,j$\rangle \in Acc_{O^n}$ and for all m such that $\langle$m,j$\rangle \in Acc_{O^n}$, $\langle$m,k$\rangle \in Acc_{O^n}$.

This assumption is needed to guarantee a coherent definition of the *local context* of a presupposition within a preliminary representation. The relevance of the constraint can be seen when we define the accessibility relation Acc_K between the sub-DRSs of a given DRS of L. We can define this relation as the transitive closure of the union of the following sets:

(a) the set of all pairs $\langle \mathrm{K}',\mathrm{K}_i \rangle$, where K_i is the i-th argument of a condition of the form $O^n(\mathrm{K}_1,...,\mathrm{K}_n)$ belonging to the condition set of K', and

(b) the set of all pairs $\langle \mathrm{K}_i,\mathrm{K}_j \rangle$, where K_i and K_j are the i-th and j-th arguments, respectively, of a condition of the form $O^n(\mathrm{K}_1,...,\mathrm{Kn})$ occurring somewhere in K and $\langle$i,j$\rangle \in Acc_{O^n}$.

(187) is easily seen to ensure that for each K' occurring in K there is a "nearest" sub-DRS K'' of K such that $\langle \mathrm{K}'',\mathrm{K}' \rangle$. The second assumption only concerns the variable binding operators. For these we assume that the discourse referent bound by the operator is accessible to each of the

[51] We assume (i) that presupposition resolution always takes place in a context which does not include information provided by the part of the sentence which contains the presupposition trigger, and (ii) that all information that is accessible from a given constituent of a preliminary representation can be used for justification of the presuppositions adjoined to that constituent. Thus the accessibility relation we need here should be asymmetric and transitive. This is a difference with the original accessibility of DRT, as defined in [Kamp, 1981a] or [Kamp and Reyle, 1993] which is transitive, antisymmetric and reflexive.

operator's arguments. More formally, suppose we write the DRS conditions fromed with the help of such an operator as "O^n_x(K$_1$,...,K$_n$)", then x is accessible from each of K$_1$, ..., K$_n$. We can express this condition in the form of a relation between DRSs, viz. by assuming that each variable binding operator O^n has an extra DRS argument of the form $\mathrm{K}_0 = \boxed{\begin{array}{c}\mathrm{x}\\ \hline \\ \\ \end{array}}$, where x is the discourse referent bound by the operator, and that the accessibility relation determined by O^n includes all pairs $\langle 0,\mathrm{i}\rangle$ (where i = 1,...,n).

It is clearly in the spirit of our general approach that the transition form DRSs to preliminary DRSs should be defined in this general way, as pertaining to any one of a large number of possible choices for L. However, some of the definitions, which are quite complex, and in our judgment are all the more difficult to understand in the abstract setting which includes an open ended class of underlying DRS languages L which are only required to obey a number of very general constraints. We therefore recommend that the reader, while he is making his way through these definitions for the first time, keeps a particular comparatively simple DRS language L$_0$ in the back of his head, where the condition forming operators are $\neg$, $\Rightarrow$, $\forall$ and ;, and where atomic conditions are of the forms x=y or P(x$_1$,...,x$_n$). For the operator set $OP_{L_0}(= \{\neg, \Rightarrow, \forall, ;\})$ the relevant information is familiar: $\forall$ is the only variable binding operator among them, all but $\neg$ are 2-place, while $\neg$ is 1-place; for $\Rightarrow$ and ; the accessibility order is $\{\langle 1,2\rangle\}$ (where for $\Rightarrow$ 1 is the restrictor and 2 the nuclear scope and for ; 1 is the first conjunct and 2 the second). For $\forall$ the accessibility relation is $\{\langle 0,1\rangle,\langle 0,2\rangle,\langle 1,2\rangle\}$, where 1 indicates the restrictor, 2 the nuclear scope and 0 the "dummy DRS" containing only the bound variable (recall the convention about the accessibility relation of bound variable operators introduced above). For the 1-place operator $\neg$ the local accessibility relation is of course $\emptyset$. Clearly this charachterisation of $Acc_{\{\neg,\Rightarrow,\forall,;\}}$ is in accordance with the constraints imposed above.

Before we proceed with the formal specifications of the various syntactic notions which we will need, a few things ought to be said about the semantics of the representation formalism we are about to define. We are facing a fundamental question here: What should we expect from a semantics for such a formalism? Different answers to this question may be possible, and different answers may be wanted on the basis of one's general view of the nature of presupposition. From our own perspective, which has informed most of what has been said in this section so far, the following answer seems adequate, and perhaps that by now the reader expected as much: The "semantics" of a preliminary representation should answer the two basic questions which are connected with it witihin a presupposition theory of the general form we have been advocating:

(i) there must be a precise model-theoretic answer to the question whether the presuppositions of the preliminary representation are verified given a global context DRS K_g; and

(ii) there must be a model-theoretic definition of whether that which remains of a preliminary representation after all its presuppositions have been justified and eliminated is true.

The second of these questions is unproblematic as long as what remains of a preliminary representation after elimination of its presuppositions is a DRS of the underlying DRS language L. For in that case this question reduces to the semantics for L, and we may assume that that has been delat with as part of the specification of L. This is what we would expect on intuitive grounds, and indeed found to be the case in all the examples we have so far considered. And as a matter of fact, it will follow from the formal definitions below that this will always be so.

What remains is the first question. At first sight this question appears daunting, because, as we have seen, the structure of preliminary representations can be very complex: Presuppositional components of a preliminary representation may have further presuppositions adjoined to them, and so on arbitrarily far down; and presuppositions can occur in the local contexts created by the operators O^n, i.e. as adjuncts to the DRSs which occur as arguments to those operators. And of course these two sources of complexity will often combine (for instance when a possessive NP such as *his rabbit* occurs within the consequent of a conditional. Presuppositions in the subordinate positions created by operators may look like they present a particularly serious problem, for in general justification of such a presupposition isn't with respect to the global context K_g as such, but with respect to the local context of the presupposition, which combines the information from K_g with information that is sentence-internal.

Nevertheless, it turns out that for the formalisms we will define question (i) has a simple solution too, which relies entirely on the model theory for the underlying language L, which is assumed to be already in place. The reason for this can be explained as follows. For any presuppositional component K' occurring somewhere in a preliminary representation K we can, given a global context K_g, define the total information available for its justification at its local context in K. Moreover, this total information at K''s local context has the form of a DRS $K_{lc(K')}$. So the question whether K' is justified given KC reduces to the question whether DRS $_{lc(K')} \models K'$ and this a question about entailment between two DRSs form L.

To determine the local context of K' in K we amalgamate all the non-presupposition parts of K which we encounter when follwing the "projection line" defined by the accessibility relation starting from the position of K''s local context all the way up to the global context K_g. Here we only collect the non-presuppositonal parts, while ignoring the presuppositions. This may seem suspicious. For a presupposition such as K′ may itself depend on other presuppositions that occur as constituents of K, for instance presuppositions that are directly left-adjoined to K′ itself. Is there anything that can be meaningfully said about the jsutification of K′, one might ask, when the justification of these other presuppositions has not yet been settled?

The answer to this question is as follows. The relation between K_g and K′ which concerns us is whether all of K's presupposiotions are justified in their respective contexts. If that is so, it will be true in particular for those presuppositional components in K which do not presuppose other such compoennts in K. This means that the information they represent is entailed by their local context (and thus by K_g together with the relevant non-presuppositional parts of K). So the question whether presuppositional components of K which depend only on presuppositions of the first ("independent") sort are justified, won't be affected by whether their presuppositions are ignored; for the information that those presuppositions represent will be part of the ontextual infomation in any case. And so on.

In other words, the analysis of presupposition justifcation we have alluded to will lead to intuitively correct answers to the question whether all presuppositions of K are justified. As soon as one presupposition is not satisfied in its local context, then our analysis can not be relied upon to give us meaningful assessments of the justification of other presuppositions of K, which depend on it. But in that case we already have a negative answer to our question in any case.

As we have seen, justification in the global discourse context of all presuppositions of the preliminary representation of a sentence is not something that can be expected. More often than not some form of accommodation will be needed. In such cases the justification anaylysis we have sketched will return a negative answer, but as things stand it will not tell us what accommodations should be made to turn the global context into one which does justify all presuppositions at once. All that the theory gives us in such cases is a criterium that decides which accommodations will be formally adequate in the sense that the resulting context does justify all presuppositions. We think it is a legitimate suspicion that this cannot be the complete story. For there are many cases where the presuppositions of a sentence that is used in a given context seem to force accommodation of a very specific kind, so much so that the accommodations feel almost like regular inferences which the discourse enables us to draw. We will see an example of this in the next subsection. For the phenomenon of "forced accommodation", where the sentence and its given context compell us to accommodate in one very specific way no explanation is given by the theory presented here.

When we argued that the entailment relation between DRSs of L is all we need to answer the question whether K_g justifies all presuppositions of K we implicitly assumed that all presuppositions were non-anaphoric. As soon as anaphoric presuppositions come into play, matters get somewhat more complicated because the non-presuppositonal parts of K may now have occurrences of discourse referents which are bound within an anaphoric presupposition on whihc the part depends. However, even this is not a real stumbling block. For justification of an anaphoric presupposition will involve linking its anaphoric discourse referents to some other discourse referent and this "antecedent" discourse referent will have to be declared in (that is, belong to the DRS-universe of) either the global context or else some non-presuppositional part of K. Suppose that there is a link for all the anaphoric discourse referents of K such that all presuppositions of K are justified given that link. Then the presuppositions can be elminated and at the same time the occurrences of the anaphoric discourse referents in nonpresuppositional parts can be replaced by their antecedents according to the link. In this way we once again obtain from K a DRS from L as definitive representation. This DRS will in general not be proper, but its free discourse referents will be declared in the global context DRS K_g, just as this would be expected on the treatment of transsentential anaphora in classical DRT.

Syntax for Preliminary Representations without Anaphoric Presuppositions

The definition of the set of preliminary representations for L is fairly straightforward except for one complication. This complication is connected with anaphoric presuppositions — those which involve anaphoric discourse referents (marked by underlining in our sample treatments in Section 4.2). We sidestep this complication for the moment by defining, as a first step, the set of preliminary representations in which anaphoric presuppositions do not occur. The definition which includes anaphoric presuppositions will follow in the next subsection. This definition is quite simple: The set PR_L^- of preliminary representations of L without anaphoric presuppositions consists of (i) the DRSs of L, and (ii) pairs of the form $\langle K,\mathrm{K}\rangle$, where K is a preliminary representation and K a set of preliminary representations (intuitively, the set of presuppositions left-adjoined to K). Moreover, complex conditions now come in two forms. On the one hand we want to admit perliminary representations of the form $\langle K,\mathrm{K}\rangle$ where K is a preliminary representation and K a set of such representations; and on the other hand we must allow for compex conditions of the form $O^n_x(\mathrm{K}_1,...,\mathrm{K}_n)$, where K_1, ..., K_n are preliminary representations. If K is of the form $\langle K,\mathrm{K}'\rangle$, then K′ may itself be again of such a form, i.e. $\langle K',\mathrm{K}''\rangle$, and so on. In a case like this both K and K' function as presuppositions for K''. We define the notion of the presupposition set in K of a quasi-DRS K′ that is a constituent of K — in symbols: PRES(K′,K) — accordingly.

(i) Suppose that K′ is a preliminary DRS that is a constituent of K and which is not part of a larger constituent $\langle K',\mathrm{K}''\rangle$. Then PRES(K′,K) $= \emptyset$.

(ii) Suppose that $\langle K,\mathrm{K}'\rangle$ is a constituent of K. Then PRES(K′,K) = PRES($\langle K,\mathrm{K}'\rangle$,K) $\cup$ K.

As usual V is the set of all discourse referents. ATCON_L is the set of atomic conditions of L.

DEFINITION 54. (Of the set PR_L^- of *Preliminary Representations* of L without anaphoric presuppositions, and the set PRCON_L^- of Conditions of such Preliminary Representations)

(i) $\mathrm{PR}_L^- ::= \langle \mathcal{P}(\mathrm{V}), \mathcal{P}(\mathrm{PRCON}_L^-)\rangle — \langle \mathcal{P}(\mathrm{PR}_L^-), \mathrm{PR}_L^-\rangle$

(ii) $\mathrm{PRCON}_L^- ::= \mathrm{ATCON}_L — \mathrm{O}^n(\mathrm{PR}_L^-,...,\mathrm{PR}_L^-)$ (with $\mathrm{O}^n \in \mathrm{OP}_L$)

($\mathcal{P}$(X) denotes the power set of X.)

In order to define the semantics for preliminary representations we need a number of notions related to the syntax of PR_L^-. These are defined under 1.-8. below.

1. We can distinguish the members of PR_L^- into two types, those preliminary representatons which are of the form $\langle K,\mathrm{K}'\rangle$ and those which are not. The latter will be called *quasi*-DRSs. (They are like DRSs in that they consist of a set of discourse referents and a set of conditions, except that the conditions need not be DRS-conditions in the strict sense of the word but can be preliminary conditions of any kind.)

2. When K $\in \mathrm{PR}_L^-$ is of the form $\langle K,\mathrm{K}'\rangle$, K′ is called the *head* of K and K the *presupposition set* of K.

3. Each member K of PR_L^- is either a quasi-DRS or it is formed from a quasi-DRS through possibly repeated adjunction of sets of presuppositions. This quasi-DRS is called the (*non-presuppositional*) *root* of K, and denoted as root(K). The definition is obvious: If K is a quasi-DRS, then root(K) = K, and if K = $\langle K,\mathrm{K}'\rangle$, then root(K) = root(K′). It follows that among the preliminary representations K′ that are part of a preliminary representation K we have: K′ is a quasi-DRS iff root(K′) = K′.

4.alt Suppose that K′ is part of a preliminary representation K (either a proper part or K itself) and that K′ is a quasi-DRS. Then the *presuppositions of* K′ *in* K, PRES(K′,K), are all those which have been added to K′ through successive left-adjunction. We define PRES(K′,K) via the auxiliary notion of a preliminary representation K″ being an Adjunction Expansion of K′ in K, ADEX(K″,K′,K). Let IMADEX(K) be the relation of immediate adjunction in K, i.e. $\langle$K″,K′$\rangle$ $\in$ IMADEX(K) iff K′, K″ are parts of K and there is a subset K of PR_L^- such that K″ $\in$ $\langle$K,K′$\rangle$. Then ADEX(K″,K′,K) iff there is a finite chain $K_1 = K', \ldots, K_m = K''$, of length $m \geq 1$, such that for $i = 1, \ldots, m-1$, $\langle K_{i+1}, K_i \rangle \in$ IMADEX(K). (Note that this entails that always ADEX(K′,K′,K).) PRES(K′,K) is defined in terms of ADEX as follows:

PRES(K′,K) =def. {K″: there are preliminary representations K_i, K_{i+1} that are constituents of K and $\mathcal{K} \subseteq PR_L^-$ such that ADEX(K_i,K′,K), $K_{i+1} = \langle \mathcal{K}, K_i \rangle$ and $K'' \in \mathcal{K}$}

4. Given a preliminary DRS K we can consider the set of all constituents of K which are presuppositions of some quasi-DRS K′ that is a constituent of K. We call this set the set of *presuppositions occurring in* K, PRES(K):

PRES(K) = {K″ | ∃ K′ (K′ is a constituent of K and K″ ∈ PRES(K′,K))}

5. Def.1 assigns to each member E of $PR_L^- \cup PRCON_L^-$ a unique parse. We can think of the parse as a decorated tree T_E, in which each node is decorated by either (i) a member of PR_L^-, (ii) a subset of PR_L^-, or (iii) a member of $PRCON_L$. Moreover, the edges of T_E are of the following types: ∈, *pres*, *head* and *Arg*$_i$*(*O^n*)*, where $O^n \in OP_L$ and $i \leq n$. Each edge connects a mother node with one of its daughters. We have an ∈-edge when either (a) the decoration of the mother node is a set of members of PR_L^- and the decoration of the daughter node is a member of that set or (b) the decoration of the mother node is a quasi-DRS and the decoration of the daughter node is one of its conditions. *pres*-edges connect a mother node decorated with a member of PR_L^- of the form $\langle \mathcal{K}, K \rangle$ with the daughter node that is decorated with the presupposition set $\mathcal{K}$ of that member; and *root*-edges connect mother nodes decorated with $\langle \mathcal{K}, K \rangle$ with the daughter node decorated with the root K. Finally, an *Arg*$_i$*(*O^n*)*-edge connects a mother node decorated with a preliminary condition of the form $O^n(K_1, \ldots, K_n)$ with the daughter node decorated with the i-th argument K_i.

Note that in the parse trees described here the discourse referents occurring in preliminary representations are ignored.

6. Each preliminary representation K can be reduced to a DRS PRESRED(K), the *presupposition reduction* of K, by eliminating all presuppositions from it. The procedure for obtaining PRESRED(K) from K may be obvious in any case, but here is a formal definition:

DEFINITION 55.

(i) PRESRED($\langle \mathcal{K}, K \rangle$) = PRESRED(K);

(ii) for a quasi-DRS K, PRESRED(K) = $\langle U_K$, {PRESRED(γ): $\gamma \in Con_K$}$\rangle$;

(iii) for γ an atomic condtion of L: PRESRED(γ) = γ

(iv) for $O^n_{(x)}$(K$_1$,...,K$_n$) $\in$ PRCON^-_L, PRESRED(γ) = $O^n_{(x)}$(PRESRED(K$_1$), ...,PRESRED(K$_n$)).

7. Let K $\in$ PR^-_L and let γ be an atomic condition occurring somewhere in K. Then there will be at least one preliminary representation K′ that is a constituent of K such that $\gamma \in$ PRESRED(K′).

The fact mentioned in 7. is the key to the definition of what it means for a preliminary representation K from PR^-_L to count as *proper*, i.e. for all the discourse referents occurring in it to be properly bound. An occurrence of a discourse referent x in some atomic condition γ which occurs somewhere in K is *bound in* K if there exists a preliminary representation K′ that is part of K such that $\gamma \in$ PRESRED(K′) and either x $\in$ U$_{K'}$ or there is a K″ in PRESRED(K) which is accessible from K′ such that x $\in$ U$_{K''}$.

DEFINITION 56. Let K $\in$ PR^-_L. K is *proper* iff for each occurrence of a discourse referent x in some atomic condition γ occurring in K there exists a K′ $\in$ PR^-_L such that K′ is a constituent of K, $\gamma \in$ PRESRED(K′) and x $\in$ U$_{K'}$.

8. Among the preliminary representations which are constituents of a given preliminary representation K, some can play a role of local context in presupposition justification. These are (i) the root of K and (ii) the root of every complex condition $\langle K$,K′$\rangle$ in K, and (iii) the roots of the arguments K$_i$ of a complex condition $O^n_{(x)}$(K$_1$,...,K$_n$) in K. We refer to this set as the set of *potential local contexts in* K and denote it as PLC(K).

All preliminary representations that are part of K and that are not in NPRP(K) are among K′s presuppositions or are part of some presupposition. We will refer to them as the *Presuppositional Representations* in K, PRESR(K).

Syntax of Preliminary Representations with Anaphoric Presuppositions

The notion of an "anaphoric presupposition", in the sense in which it was used in Section 4.2, involves that of "anaphoric" discourse referents, discourse referents which must, as part of the presupposition's justification, find antecedents in some accessible context. In the examples of anaphoric presuppositions we have seen there was never more than one anaphoric discourse referent per anaphoric presupposition, but this is a restriction that we cannot expect to hold generally. So we want to allow for arbitrary sets of anaphoric discourse referents. In other words, the set of anaphoric discourse referents will in general be some subset of the main Universe U of such a presupposition representation. We allow any subset between $\emptyset$ and U inclusive, the case of being the non-anaphoric — or "purely propositional" — presuppositions being that where the set = $\emptyset$.

The simplest way to formalise this notion of an anaphoric presuppositon representation is to replace in our definitions of PR_L and PRCON_L the DRS universes U everywhere by pairs $\langle$U,A$\rangle$, with A $\subseteq$ U. A is the set of anaphoric discourse referents of the given representation with "universe" $\langle$U,A$\rangle$.

A slight further complication is that anaphoric discourse referents have no business in the non-presuppositional parts of representations. That is, if $\langle\langle$U,A$\rangle$, Con$\rangle \in$ PLC(K) for some K, then A should be $\emptyset$. We denote the set of preliminary representations K for which this condition holds as PR'_L. This more restricted set also is now the resource from which complex conditions are built. The need to distinguish between PR_L and PR'_L entails that we now need a definition by simultaneous recursion of the three sets PR_L, PR'_L and PRCON_L. It is convenient in this connection to deviate a little more from the strict Backus-Naur format than we did in Def. 54.

DEFINITION 57. (Of the set PR_L of *Preliminary Representations* of L with anaphoric presuppositions, and the set PRCON_L of Conditions of such Preliminary Representations)

$\mathrm{PR}_L ::= \langle\langle \mathrm{U,A}\rangle, \mathcal{P}(\mathrm{PRCON}_L)\rangle$, where $\mathrm{U} \in \mathcal{P}(\mathrm{V})$ and $\mathrm{A} \subseteq \mathrm{U}$
$\qquad — \langle \mathcal{P}(\mathrm{PR}_L), \mathrm{PR}_L\rangle$

$\mathrm{PR}'_L ::= \langle\langle \mathrm{U},\emptyset\rangle, \mathcal{P}(\mathrm{PRCON}_L)\rangle$, where $\mathrm{U} \in \mathrm{P}(\mathrm{V})$
$\qquad — \langle \mathcal{P}(\mathrm{PR}_L), \mathrm{PR}'_L\rangle$

$\mathrm{PRCON}_L ::= \mathrm{ATCON}_L — O^n(\mathrm{PR}'_L,\ldots,\mathrm{PR}'_L)$ (with $O^n \in OP_L$)

N.B. there is a one-one correspondance between the preliminary representations given in Def. 54 and those preliminary representations in the sense of Def. 57 in which all universes are of the form $\langle$U,$\emptyset\rangle$. Let us denote the subset of these preliminary representations in the sense of Def. 57 which correspond in this way to members of PR^-_L as PR^{-1}_L. Then we evidently have $\mathrm{PR}^{-1}_L \subset \mathrm{PR}'_L \subset \mathrm{PR}_L$.

All notions defined in the last section for members of PR^-_L generalise straightforwardly to the sets PR_L and PR'_L. In particular, every member of PR_L has a unique parse, which can be represented by a parse tree of the same form as defined on page 298. The only exception is the notion of a proper representation. This notion requires renewed attention because the anaphoric presuppositions create situations in which an occurrence of a discourse referent x in some atomic condition belonging to a part representation K′ is bound by the occurrence of x in the universe of some presupposition of K′. We have seen several instances of this in Section 4.2. For example, in (153) the occurrence of u in the condition "white(u)" of the non-presuppositional part of the representation is bound by the occurrence of $\underline{\mathrm{u}}$ in the universe of the presupposition left-adjoined to this part.

We will assume that presuppositional binding of discourse referents is always of this comparatively simple form: If an occurrence of x in some atomic condition belonging to some quasi-DRS K′ in a preliminary representation is bound presuppositionally, then this can be only through the presence of x in the set A of anaphoric discourse referents of a preliminary representation in PRES(K′,K). (This entails in particular that if x is any non-anaphoric discourse referent belonging to the universe of a presupposition K″ of K′ (i.e. if this universe is $\langle$U,A$\rangle$, then x $\in$ U$\backslash$A), then x will not occur in atomic conditions belonging to K′. For an illustration, see the discourse referent x for the possessive pronoun *his* in (150).)

These assumptions lead us to the notion of the extended universe of a quasi-DRS K′ belonging to some preliminary representation K. We denote this set as EU(K′,K), and sometimes, when it is clear which K is at issue, as $\mathrm{EU}_{K'}$. EU(K′,K) consists of the universe of K′ itself together with the sets of anaphoric discourse referents of all members of the presupposition set of K′ in K. In other words:

$$\begin{array}{l}\mathrm{EU}(\mathrm{K}',\mathrm{K}) = \\ \mathrm{U}_{K'} \cup \bigcup \{\mathrm{A}\colon (\exists\, \mathrm{K}'',\mathrm{U},\mathrm{Con})(\mathrm{K}'' \in \mathrm{PRES}(\mathrm{K}',\mathrm{K})\ \&\ \mathrm{K}'' = \langle\langle \mathrm{U},\mathrm{A}\rangle,\mathrm{Con}\rangle\}\end{array}$$

$\mathrm{EU}_{K'}$ replaces $\mathrm{U}_{K'}$ in a couple of the auxiliary notions introduced above. First, the definition of the reduction PRESRED now has to be modified in that if K′ is a quasi-DRS, then

$$\mathrm{PRESRED}(\mathrm{K}') = \langle \mathrm{EU}_{K'}, \{\mathrm{PRESRED}(\gamma) : \gamma \in \mathrm{Con}_{K'}\}\rangle.$$

(This renders the definition on the larger preliminary representation K of which K′ is considered a part, so that PRESRED now becomes dependent on this second parameter as well. Thus, strictly speaking the definition is now of a 2-place function PRESRED(K′,K). But we will persist with the earlier notation and only mention the argument K′. Secondly, need to adapt the definition of what it is for a preliminary representation to be proper.

DEFINITION 58. Let K $\in$ PR_L. K is *proper* iff for each occurrence of a discourse referent x in some atomic condition γ occurring in K there exists a K′ $\in$ PR_L such that K′ is a constituent of K, $\gamma \in$ PRESRED(K′) and x $\in$ $\mathrm{EU}_{K'}$.

Local Contexts

Suppose that K is a member of PR'_L and that K′ is a presupposition occurring in K, i.e. K′ $\in$ PRES(K). Justification of K′ takes place in the *local context* of K′ in K whenever possible, and only if K′ has no local context in K in the *global context.* The local context of K is intuitively the first quasi-DRS K″ in K (if any) which one encounters going up the parse tree T_K of K, starting from K′. If such a quasi-DRS K″ is reached , this will always mean that the last edge of the path running from K$'$ to K$''$ is an $\in$-edge and taht if m and d are the mother node and daughter node this edge connects, then the decoration of d is a condition belonging to the condition set of the duration of m. There then are two possibilities: (i) the condition at d is of the form $\langle K,\mathrm{K}'''\rangle$ with $K \in \mathcal{P}(\mathrm{PR}_L)$, $\mathrm{K}''' \in \mathrm{PR}_L$; (ii) the condition is of the form $O^n_{(x)}(\mathrm{K}_1,...,\mathrm{K}_n)$. In the first case K$''$ is the local context of K$'$ in K. The second case is a little more complicated. In this case the node will itself be the mother node of an $\mathrm{Arg}_i(O^n)$-edge along the given path, and the corresponding daughter d$'$ will be decorated with K_i. If there is a j such that $\langle$j,i$\rangle \in \mathrm{Acc}_{O^n}$ then the local context of K$'$ will be the root of that K_h (h$\neq$i, h$\leq$n) such that $\langle$h,i$\rangle \in \mathrm{Acc}_{O^n}$ and for all j such that $\langle$j,i$\rangle \in \mathrm{Acc}_{O^n}$ $\langle$j,h$\rangle \in \mathrm{Acc}_{O^n}$. (I.e. the root of K_h which is the nearest to K of the arguments of O^n which are accessible from K_i.) If for no j $\leq$ n $\langle$j,i$\rangle \in \mathrm{Acc}_{O^n}$, then the local context of K_i is K$''$.

We can define this notion of local context formally on the basis of a notion of accessibility for preliminary representations which we define first.

DEFINITION 59. Let K be a preliminary representation, then the *accessibility relation on* K, Acc_K, is the set of all pairs $\langle$K$''$,K$'\rangle$, where K$''$ and K$'$ are constitutents of K, K$'$ is a preliminary representation and K$''$ is a quasi-DRS, is defined as the transitive closure of the relation Acc^0_K. Acc^0_K consists (i) of all pairs $\langle$ root(K_j), $\mathrm{K}_i\rangle$ such that $O^n_{(x)}(\mathrm{K}_1,...,\mathrm{K}_n)$ occurs in K and $\langle$j,i$\rangle \in \mathrm{Acc}_{O^n}$; and (ii) of all pairs $\langle$ K$''$, K$'\rangle$ satisfying the following condition:

(a) (which is close to the one already informally described)
K$'$ is a preliminary representation that is a constituent of K and K$''$ is determined as follows: go up through the construction tree T_K of K, starting from K$'$. K$''$ is the decoration of the first node along this path whose decoration is a quasi-DRS.

Like the accessibility relation between sub-DRSs of a given DRS, Acc_K is a strict partial order. Furthermore it is not hard to verify that if K$' \in$ Ran(Acc_K) (i.e. there are "sentence-internal" contexts of K$'$), then there is a "nearest" such K$''$, i.e. $\langle$K$''$,K$'\rangle \in \mathrm{Acc}_K$ and for all $\langle$K$'''$,K$'\rangle \in \mathrm{Acc}_K$ either K$'''$ = K$''$ or$\langle$K$'''$,K$''\rangle \in \mathrm{Acc}_K$. And, finally, whenever $\langle$K$''$,K$'\rangle \in \mathrm{Acc}_K$, then K$''$ is a quasi-DRS and K$'' \in$ PLC(K).

We are now in a position to define the local context of a presuppositional component K$'$ of a preliminary representation K. There are in fact three related but distinct notions of local context that we will need. The first is the one which we have informally described already: The local context of K$'$ in K in this sense is that K$''$ in PLC(K) which is nearest to K$'$ in the sense of Acc_K (provided any such K$''$ exists; if not, then K$'$ doesn't have a local context). We represent this notion of local context in the form of a 3-place relation between K, K′ and its local context K″ and denote the relation as "LocConK(K″,K′,K)").

The second notion is that of the *Sentence-Internal Information available for presupposition justification at the local context of* K′ *in* K, which we denote as SILC(K′). Intuitively SILC(K′,K) consists of the totality of presupposition-free information that is available at all "sentence-internal" contexts accessible from K′, i.e. all quasi-DRSs K″ such that $\langle$K″,K′$\rangle \in \mathrm{Acc}_K$. All local contexts, we saw, are quasi-DRSs. But what is the "presupposition-free" information of a quasi-DRS? The definition is pretty much as the term suggests: the *presupposition-free information* of a quasi-DRS K″ consists of the discourse referents of K″ together with those conditions which contain no presuppositions, and thus are DRS-conditions of the language L.

Let K be a quasi-DRS from PR_L, then
$\mathrm{PF(K)} = \langle\ \mathrm{U}_K\ ,\ \{\ \gamma \in \mathrm{Con}_K : \gamma \in \mathrm{CON}_L\ \}\ \rangle$

We can now define SILC(K′,K) as the merge of all the DRSs PF(K″) for $\langle$K″,K′$\rangle \in \mathrm{Acc}_K$.

DEFINITION 60. Let K‴ be a preliminary representation that is a constituent of a preliminary representation K. Then
$\mathrm{SILC(K',K)} = \biguplus\ \{\ \mathrm{PF(K'')} : \langle \mathrm{K'',K'} \rangle \in \mathrm{Acc}_K\}$.

(Here $\biguplus$ represents the merge of a set of DRSs. See the end of Section 3.2.) NB. In case K′ has no local context in K, then the argument set of $\biguplus$ in the definition above is empty and $\mathrm{SILC(K',K)} = \biguplus\emptyset = \langle\emptyset,\emptyset\rangle$ (the empty DRS).

The third notion of local context is very close to the second. This is the *total information available for presupposition justification at the local context of* K′ *in* K, TILC(K′,K). TILC(K′,K) is the merge of SILC(K′,K) with the global context DRS K_g.

Def. 61 repeats the definitions of the three notions.

DEFINITION 61. (Local Context of K′ in K; *Total Information Available at the Local Context* of K′ in K; and *Sentence-Internal Information available at the Local Context* of K′ in K)

Let K be a preliminary representation and K_g a DRS. (Intuitively, K_g represents the context in which the sentence represented by K is uttered.) Let K′, K″ be preliminary representations that are constituents of K.

(i) LocCon_K(K″,K′,K) iff

  (a) $\langle$K″, K′$\rangle \in \mathrm{Acc}_K$ and

  (b) for all K‴ such that $\langle$K‴,K′$\rangle \in \mathrm{Acc}_K$, $\langle$K‴,K″$\rangle \in \mathrm{Acc}_K$.

(ii) $\mathrm{SILC}(\mathrm{K}',\mathrm{K}) = \uplus \{\mathrm{PF}(\mathrm{K}''): \langle \mathrm{K}'', \mathrm{K}' \rangle \in \mathrm{Acc}_K\}$

(iii) $\mathrm{TILC}(\mathrm{K}',\mathrm{K},\mathrm{K}_g) = \mathrm{SILC}(\mathrm{K}') \cup \mathrm{K}_g$

The point of these different notions is as follows. Assume that $\mathrm{K}' \in \mathrm{PRES}(\mathrm{K})$. The local context K″ of K′ in K is intuitively the lowest point in the logical structure of K where justification of the presupposition K′ is possible. We assume that a presupposition is always justified at this lowest possible point. In other words — this is one respect in which the present proposal differs from other DRT-based accounts:

(188) Presupposition justification always takes place at the local context.

However, the contextual information that is available for justification of K′ at its local context K″ includes not only the information contained in K″ itself but also that contained in all K‴ in K which are accessible from K″ (and thus from K′) as well as that of the global context K_g. (Thus the more local a context, the more information it makes available.) It follows from this stipulation that if justification of K′ can succeed at all, it will succeed at K″. So the assumption (188) that presuppositions are always justified at their local context isn't shouldn't be seen as an empirical claim. It only reflects a particular perspective on the nature of presupposition justification.

Sometimes justification of K′ at its local context K″ is possible on the basis of SILC(K′,K) alone. These are the cases which the classical presupposition literature describes as instances of "local satisfaction", or "local binding",[52] cases where the presupposition, being justifiable without any appeal to K_g, disappears as a presupposition of the full sentence which contains its trigger — in other words, where the presupposition "doesn't project". It disappears because the constraints it imposes on context are satisfied in any case. Thus, as far as it is concerned, the sentence could be uttered in any global context.

Semantics for Preliminary Representations

As explained above, the "semantics" of preliminary DRSs as we understand it only concerns the question whether the presuppositions of a preliminary DRS K are justified in a global context K_g. And this question, we already saw, has a positive answer iff for every $\mathrm{K}' \in \mathrm{PRES}(\mathrm{K})$ K′ is entailed by the total information at its local context. In the case where none of the presuppositions of K are anaphoric this amounts simply to: $\mathrm{TILC}(\mathrm{K}',\mathrm{K},\mathrm{K}_g) \models \mathrm{K}'$ for all $\mathrm{K}' \in \mathrm{PRES}(\mathrm{K})$.

[52] Or alternatively, as "intermediate satisfaction" or "binding". Note that the use we make of "local" corresponds to what others have called "intermediate" (cf. in particular [van der Sandt, 1992]).

That is all that needs to be said for this case. If the answer is positive, then K can be reduced to the DRS PRESRED(K). The questions of truth and verification for such DRSs are a matter for the semantics of the underlying language L, as is the definition of TILC(K′,K,K_g) $\models$ K′. In case K contains anaphoric presuppositions the matter is more complicated. Justification of the presuppositions of K must now be made dependend on a resolution of the anaphoric discourse referents, and we need to spend some care on the definition of what a possible resolution is for the anaphoric discourse referents of K given a global context K_g.

A possible resolution of K (given K_g) must link each anaphoric discourse referent u occurring in K with a possible antecedent x. For x to be a possible antecedent for u, x must (i) accessible from the position of u in the sense familiar from standard DRT, and (ii) x must belong to the universe of a quasi-DRS K″ which qualifies as context from the perspective of the anaphoric presupposition K′ which contains u as a member of its universe. Both these requirements are fulfilled if $\langle K'', K' \rangle \in Acc_K$. Let us make these considerations more explicit. Suppose that K $\in$ PR_L and that x is a discourse referent which belongs to the set of anaphoric discourse referents $A_{K'}$ for some constituent K′ of PRES(K) (in other words $U_{K'} = \langle U, A \rangle$). The set of potential antecedents for x is then the union of all universes of quasi-DRSs K″ in K such that $\langle K'', K' \rangle \in Acc_K$ together with the universe of the context DRS K_g. We distinguish between the set IPA(x,K) of those potential antecedents of x which are "internal to" K, and the total set of potential antecedents, given the context K_g, PA(x,K,K_g). Formally:

$$\mathrm{IPA}(x,K) = \bigcup \{U_{K''}\colon \langle K'', K' \rangle \in Acc_K\}$$
$$\mathrm{PA}(x,K,K_g) = \mathrm{IPA}(x,K) \cup U_{K_g}$$

IPA(x,K) and PA(x,K,K_g) enable us to define the notion of a potential resolution of the anaphoric discourse referents of K:

> A *potential anaphoric resolution for* a preliminary representation K belonging to PR_L, *given the global context* K_g, is a function r from discourse referents to discourse referents whose domain consists of the anaphoric discourse referents occurring in K and which is such that for any such discourse referent x, r(x) $\in$ PA(x,K,K_g).
>
> We say that r *resolves x sentence-internally* iff r(x) $\in$ IPA(x,K).

Exactly how the anaphoric discourse referents of a preliminary DRS K should be resolved in a context DRS K_g — i.e. which of the possible resolutions should be chosen — is a problem which classical DRT made it a policy to leave to other theories. We will adopt this policy here too. That is, we consider, given K and K_g, any one of the possible resolutions r for the anaphoric discourse referents of K, given K_g, and then consider whether for each of the presuppositions K′ of K TILC(K′,K,K_g) $\models$ K′ given this choice of r. We abbreviate this relation as TILC(K′,K,K_g) $\models_r$ K′. This relation holds provided the discourse referents of K′ are always assigned the same values as the discourse referents in TILC(K′,K,K_g) to which r resolves them. In other words, for any DRSs K_1, K_2 $K_1 \models_r K_2$ iff for any model $\mathcal{M}$, world w of $\mathcal{M}$ and time t of $\mathcal{M}$ r-verification of K_1 within w at t, entails r-verification of K_2 within w at t, where r-verification is defined in the same way as verification except that the embedding functions f involved must all satisfy the following condition:

$$\text{if x} \in \text{Dom}(r) \cup \text{Dom}(f)\text{, then } r(\text{x}) \in \text{Dom}(f) \text{ and } f(\text{x}) = f(r(\text{x})).$$

Suppose we can find a resolution r for K, given K_g, such that all presuppositions are justified by K_g. Then, as for the case where K contains no anaphoric presuppositions K should be reducible to a DRS K′ by eliminating all presuppositions from it. We must now take care, however, that when an anaphoric presupposition is removed, and with it an anaphoric discourse referent or discourse referents occurring in its universe, then the occurrences of the anaphoric discourse referents in conditions belonging to the non-presuppositional parts of K, which are not removed should be replaced by their antecedents under r. We obtain the desired result by reducing K first through application of the operator PRESRED and then replacing discourse referents in PRESRED(K) which also occur in Dom(r) everywhere by their r-values. The result will be a DRS r(PRESRED(K)) which need not be proper, but where free discourse referents will belong to the universe of K_g.

4.4 Accommodation and Inference.

The semantics developed in Section 4.3 tells us when a sentence, uttered in a context K_g and preliminarily represented as K, is true or false in a model. Part of what it tells us is that the sentence will be either true or false only if all its presuppositions are justified. The examples we have looked at in Section 4.2 have given us a taste of how stringent this requirement is. It is normal for a sentence to generate presuppositions, and usually not just one but a whole bunch of them: Joint satisfaction of all those presuppositions is a constraint that it is in general not easy for utterance contexts to meet. This implies that if a speaker wants to make an assertion that is true, he will have to proceed with great caution in general, lest this sentence generate a presupposition that in the given context isn't warranted.

When we see how language is actually used and interpreted, this conclusion appears alarmist. Far fewer utterances seem inappropriate than it predicts. The reason for this discrepancy is that human interpreters are generous accommodators. Many presuppositions are accommodated quasi-automatically by interpreters who don't seem to be aware of the fact that they are doing so. Normally it is only when an accommodation that is needed goes against something that the interpreter believes that he will be concious of what he is doing – that he is adjusting his assumptions in such a way that the utterance makes sense against their backgound.

When an accommodation which the interpreter perceives as required is in conflict with what he believes, he may nevertheless make it and revise his beliefs accordingly. But if these beliefs are too amply entrenched — when he is quite certain of them, then he will refuse to make the accommodation, and in such cases, we already stated in the last section, we regard the interpretation process as breaking down: a perliminary representation for the utterance can be constructed, but there is no way (not even one involving accommodation) to integrate it into the context. How easy a presuppositional constraint ca be accommodated seems to vary between presuppositions, and in particular as a function of their triggers. Pronouns are a notorious case in point (if they are included among the presupposition triggers at all, as we have done here). Factive presuppositions are among the kinds of presuppositions that are accomodated with great ease. If A tells B that Fred is relaxed that his proposal wasn't accepted, and B didn't know that Fred proposed, B will assume this almost as a matter of course. (He will .... only if he is convinced that Fred didn't propose.) Such presuppositions differ from presuppositions triggered by a word like *too*. An utterance containing *too* (or an equivalent expression such as *also*, *as well* and some others) gives rise to a presupposition that should be justifiable in the context established by the immediately prededing discourse. These "anaphoric" presuppositions (in the sense of "anaphoric" used by Kripke which is different from the sense in which we have been using the term) are hard to accommodate because the discourse context is as accessible to the interpreter as it is to the speaker: It is constituted by that what has just been said and that is equally known to both parties.[53]

The distinction between presuppositions that are "anaphoric" in the present sense and those that are not, is only one among a number which we expect a detailed theory of presupposition will have to draw.[54] And it

[53]The exception that confirms the rule is where a hearer drops in on an ongoing conversation and the first sentence he hears is one containing *too*. In these circumstances a *too*-presupposition will be readily accommodated, but not only in the sense of being entailed by something that is assumed to be true but as something that was actually *said* in virtue of which the presupposition is justified.

[54]Another distinction has to do with how easily a presupposition is cancellable under negation. Cancellation of the existence presupposition carried by a definite description, while possible, requires genuine effort — you can say *The exhibition wasn't opened by the King of France.* but something like the *because*-clause is indispensible lest the main clause be misunderstood. The matter is different for the pre-state presuppositions of change-of-state verbs. Take for instance the transitive verb *open*. You can open something at a time t only when at t that thing isn't open yet — y's being closed is a neccessary pre-state of an event of opening y. And this condition appears to be presuppositional insofar as there is a tendency to interpret negated statements like *He didn't open the*

isn't clear at this point which of these distinctions are binary and which a matter of graduation. The need to draw such distinctions between types of presuppositions exists in particular for a theory like the one sketched here, which is very liberal in what it includes among the range of phenomena to which it applies.

As far as we can see the DR-theoretical bases of the present theory is of little help in telling what these distinctions are, and we have more to say on their account here. There is another issue connected with accommodation, however, where the DRT-approach outlined here does appear to be of use. Often presupposition accommodation strikes us not only as possible, but in fact as necessary. In such cases the accommodation seems to be forced upon the interpreter, and the accommodated information seems more like an inference from the uttered sentence or discourse than like an assumption which the interpreter chooses to make for the sake of restoring coherence.

We present two examples of this phenomenon which have been discussed in the literature. In these examples the inferential effect of accommodation seems particularly compelling. One of them, given in (189.b), is a sentence that was first presented by Kripke in a lecture that is often cited, but of which no canonical textual version seems available.[55]

(189) a. We shouldn't have pizza on John's birthday, if we are going to have pizza on Mary's birthday.

b. We shouldn't have pizza again on John's birthday, if we are going to have pizza on Mary's birthday.

(189.b) strongly invites the inference that Mary's birthday is before John's birthday. (189.a) does not seem to carry this implication — if at all, then surely much less forcefully than (189.b). The difference can only be the presence in (189.b) of *again*.

What is the explanation of this effect? As we saw in 2.3, occurrences of *again* trigger presuppositions to the effect that an event of the same type as that described in the clause containing the occurrence happened before the described event. In other words, we have the presupposition that at some time before the event of "we" having pizza at John's birthday there was another event of "we" having pizza. In (189.b) this presupposition is generated within the consequent of the conditional, so its local context is the antecedent of the conditional. As it stands, the information contained by the antecedent goes a fair way towards justifying the presupposition, since the antecedent does speak of a pizza eating event with "we" as agent. But it

window. as implying that the window was closed at the time in question. Nevertheless when someone asks you: *Did you open the window while I was out of the room?* you can quite legitimately answer with a simle *no* even if as a matter of fact the window had been open all day long.

[55] fn on Kripke's presupposition lecture.

doesn't stretch all the way. What is still missing is the information that the event spoken of in the antecedent temporally precedes the one spoken of in the consequent. Still, the antecedent gets us so close to a justification of the *again*-presupposition that the impression that is meant to be understood as the justification of the presupposition seems virtually inescapable. So the recipient of (189.b), who is uninformed about the dates of the two birthdays, will feel impelled – he will *conclude* — that Mary's birthday comes before John's.

That this "conclusion" is mediated by presupposition justification finds further support in the circumstance that when (189.b) is offered as follow-up to a sentence which talks about yet another pizza eating event the conclusion may be blocked. Thus consider (190)

(190) We have just had pizza on Billie's birthday. So we shouldn't have pizza again on John's birthday, if we have pizza on Mary's birthday.

In (190) the *again*-presupposition can be justified in the context provided by the first sentence, (Indeed, since the first sentence speaks of such an event in the past and the main clause of the second sentence of one in the future, justification doesn't need accommodation in this case.) Since the presupposition can be justified in this way, there is no need to use the information of the *if*-clause for this purpose, so there is no need to accommodate that Mary's birthday precedes that of John.

In our second example the inferential flavour of accommodation is even stronger, and it is hard to see how it could be suspended by providing more context. This example is a three-sentence discourse, given in (191).

(191) I gave the workmen a generous tip. One went out of his way to thank me. The other one left without saying a word.

Anyone who reads these two lines knows that the number of workmen to whom the speaker gave a generous tip must have been two. How does this knowledge arise? It clearly depends on the subject phrases of the three sentences. The dependence on the subject of the third sentence, for instance, becomes visible when we replace it by certain alternatives, while leaving everything else the same. Thus, for each of the substitutions *Another one*, *The other two*, *Two others*, *One*, *Two* for *The other one* the conclusion about how many workmen there are will be different. Likewise, dependence on the subject of the second sentence is shown by replacing its subject by, e.g., *Two*, *At least two* or *Another one*.

The reason why the subject phrases of (191) produce the inferential effect observed has to do with their specific anaphoric properties — or in other words (given our liberal use of the term "presupposition") with the specific presuppositions to which these phrases give rise. To give an idea of the interpretational mechanisms that are involved in this case without enmeshing us into too much detail, we will focus on the third subject NP *The other one*. We assume that the phrase *the workmen* has introduced a discourse referent X standing for a set of two or more workmen, and that the NP *one* has been interpreted as introducing a discourse referent y together with the condition "y $\in$ X" which says that the individual represented by y is one of the members of X.

When we look at the NP *the other one* more closely, we see that it gives rise to a "bundle" of presuppositions, each one of which is connected either with the lexical meaning of one of the words of which the NP is made up or else with a morphological feature. They are:

(i) an anaphoric presupposition triggered by *one*, to the effect that there is a (nominal) predicate in context with which *one* can be identified

(ii) a doubly anaphoric presupposition triggered by *other*, to the effect that the referent of the NP is distinct from some other individual or individuals of the same type, or belonging to the same set.

This presupposition is doubly anaphoric in that both of the following items must be identifiable in context:

(a) the type or set which contains both the referent of the NP and the individual or individuals from which it differs, and

(b) the other individual or individuals belonging to this type or set

(iii) A presupposition connected with the fact that the NP is in the singular, to the effect that the NP's referent is a single individual, rather than a set of two or more individuals.

(iv) A maximality presupposition connected with the definite article *the*, to the effect that in the given context the referent of the phrase exhausts the extension of the descriptive content of the NP.[56]

The set X and its member y, both of which are part of the context within which the third sentence must be interpreted, provide a very good basis for satisfying this complex of presuppositions. Let's assume that ζ is the discourse referent introduced to represent the referent of the NP *The other one*. Identifying the predicate "$\in$ X" with *one*, and y with the presupposed individual(s) falling under the relevant predicate, which after this first identification becomes "$\in$ X", deals with the presuppositions (i) and (ii), triggered by *one* and *other*. What remains is the singularity presupposition (iii), which says that ζ represents an individual, and the maximality

[56]The existence and uniqueness presuppositions which we assumed for definite descriptions in 4.2 can be seen as a combination of these three factors: (i) existence, (ii) maximality, and (iii) cardinality 1. Here these factors are attributed to (i) the very fact that an NP contributes a discourse referent, (ii) *the*, and (iii) the singular.

presupposition (iv), to the effect that this individual exhausts the predicate of being a member of X that is distinct from y. Accommodating the assumption that the cardinality of X is 2, both these remaining presuppositions are fulfilled as well. It is also clear that no other assumption about the cardinality of X will lead to justification of both of these presuppositions.

As in the case of (189.b), the interpreter is compelled to accommodate this information. (In fact, the accommodation comes so naturally to the human interpreter that audiences to which (191) is presented tend to have considerable difficulty at first in seeing what the point of the example could be.) Moreover, we do not see any way of embedding (191) in a larger discourse in which the "inference" is cancelled — in this regard (191) appears to differ from (189.b).

Both examples suggest that our strategies for dealing with presuppositions in discourse involve some kind of "economy principle", which forces the interpreter to choose that presupposition resolution which gets by with the smallest amount of additional (i.e. accommodated) information. The extra information which must still be accommodated even when this most "economical" solution is chosen then appears as something that the discourse entails.[57]

One intuitively attractive way of thinking about accommodation for presupposition justification is to see it as a special form of *abduction*:[58] The interpreter of an utterance is trying to find the "simplest" explanation for why the speaker would have uttered an expression which generates those presuppositions that his actual utterance does generate. From this abduction-theoretic perspective the fact that accommodations are often so compelling as to look like inference. One "abductive" accommodation may, when compared with possible alternatives, appear so unequivocally superior that the interpreter simply has no choice but to adopt it as the correct way to justify the utterance. Hence the impression that the accommodation is entailed by the sentence or discourse for which it is needed.

But a caveat is in order. The abduction-theoretic pespective allows us to see presupposition accommodation as part of a much more general type of problem — that of coming up with hypotheses which account for observations which would otherwise remain unexplained. However, the mechanisms of presupposition accommodation and the contraints to which it is answerable are closely adapted to the special structure of language and the principles of verbal communication, and so we can't expect that seeing

[57]For some discussion of this aspect of presupposition justification, as well as for a motivation of the term "justification" which we have used freely within this Section, see [Kamp and Roßdeutscher, 1994] [DRS-Construction and Lexically Driven Inference, Theoretical Linguistics Vol (20, nr. 2/3, pp. 165-235].

[58]The abduction-theoretic perspective on presupposition accommodation is argued persuasively and worked out in considerable detail in the doctoral dissertation of Krause. See [Krause, 2001].

accommodation as a form of abduction will go very far in helping us to determine its special properties. All the hard work that is needed to uncover those mechanisms and contraints remains, even if the abduction-theoretic perspective promises to give us a plausible way of interpreting the results once we have them in front of us.

Our final point in this Section concerns accommodations that are needed to justify presuppositions occurring in embedded positions, such as, e.g. the *again*-presupposition of (189.b) and (190). So far, we have said nothing about *where* such accommodations are made: Is the accommodated information added to the global context, to the local context of the presupposition whose justification requires it, or at some context "intermediate" between these two in cases where there are such intermediate contexts. (In the case of (189.b) there aren't any intermediate contexts, but often there are.)

In the case of (189.b) the question seems to be only of formal interest, for the proposition that Mary has her birthday before John is true or false categorically — its truth is not dependent on whether "we" have pizza on Mary's birthday. In other words, the accommodation is one which affects the global context; whether we enter the accommodated information into the representation of the global context itself or into that of the antecedent of the conditional doesn't make a real difference one way or the other.

This is not so, however, for a sentence like (192).

(192) Every Angelino uses his car to go to work; most New Yorkers use it only during the weekend.[59]

When someone is offered this sentence out of the blue, the question whether all people from LA or New Yorkers have a car or his knowledge that many New Yorkers don't are unlikely to bother him. He will assume that the speaker intends to speak only of those people from Los Angeles and New York who do have cars. That is, he will interpret (192) as equivalent to (193).

(193) Every Angelino who has a car uses his car to go to work; most New Yorkers who have a car use it only during the weekend.

This observation has sometimes been taken as evidence that in some instances accommodation takes place at a non-global level. In relation to the first sentence of (192) the argument is as follows. The definite description *his car* creates a presupposition to the effect that the relevant individual y has a car, and the pronoun *his* contained in it gives rise to a further presupposition involving some anaphoric discourse referent u. These presuppositions are generated within the nuclear scope of the universal quantifier over Angelinos, so their local context (in our sense of "local") is the

[59] Cf. [Beaver, 2004]

restrictor of this quantifier. Following the representation in (150), we get an additional presupposition for the contextual restrictor C of the existence-and-uniqueness presupposition of the definite description. (194) gives the preliminary representation.

(194)
$$\boxed{\begin{array}{|c|}\hline \mathrm{x} \\ \hline \mathrm{Angelino(x)} \\ \hline\end{array} \;\; \left\langle \begin{array}{c}\forall\\ \mathrm{x}\end{array} \right\rangle \;\; \left\langle \mathrm{K}, \begin{array}{|c|}\hline \mathrm{v} \\ \hline \mathrm{ttgtw(x,v)} \\ \hline\end{array} \right\rangle}$$

$$\mathrm{K} = \left\langle \left\{ \begin{array}{|c|}\hline \underline{\mathrm{y}} \\ \hline \mathrm{male(y)} \\ \mathrm{pers(y)} \\ \hline\end{array} \right\}, \left\langle \left\{ \begin{array}{|c|}\hline \underline{\mathrm{C}}\ \mathrm{c} \\ \hline \mathrm{car(c)} \\ \mathrm{C(c)} \\ \mathrm{have(y,c)} \\ \hline\end{array} \right\}, \begin{array}{|c|}\hline \mathrm{u}^{=1} \\ \hline \mathrm{car(u)} \\ \mathrm{C(u)} \\ \mathrm{have(y,u)} \\ \hline\end{array}) \right\rangle \right\rangle$$

If we accommodate the three presuppositions adjoined to the nuclar scope of the quantifier in (194), we get the interpretation of the first conjunct of (192) that is given by the paraphrase in (193). And it seems that that is the only way in which we can obtain this reading. Thus, the argument goes, non-global accommodation is sometimes needed.

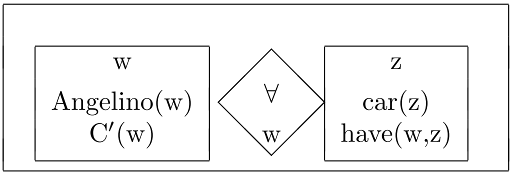
As Beaver [Beaver, 1997] has observed, the problem with non-global accommodation is that it easily leads to overgeneration – that is, of readings for sentences with embedded presuppositions which human interpreters do not get. Moreover, we believe it to be in the spirit of his general view of presupposition accommodation to maintain that even in a case like (192) accommodation is a matter of adjusting the global context. The reason why global accommodation can give us the desired reading in the case of the first sentence of (192) is connected with an omission in its preliminary representation given in (194). Note that by our own standards (194) is incomplete. It fails to represent the contextual restrictor which, we already noted in 4.2, enters into the interpretation of quantifiers no less than in the (quantificational) uniqueness presuppositions of definite descriptions. When we add a representation of this restrictor and its representation to (194), as in (195), then the reading we are after can be obtained by global restriction too. In (195) it is assumed that the extension of the quantification restrictor C′ contains at least one object of the kind explicitly specified by the quantifying NP.

(195)
$$\boxed{\boxed{\left\langle \left\{ \begin{array}{|c|}\hline \underline{\mathrm{C'}}\ \mathrm{x'} \\ \hline \mathrm{Angelino(x')} \\ \mathrm{C'(x')} \\ \hline\end{array} \right\}, \begin{array}{|c|}\hline \mathrm{x} \\ \hline \mathrm{Angelino(x} \\ \mathrm{C'(x)} \\ \hline\end{array} \right\rangle} \;\; \left\langle \begin{array}{c}\forall\\ \mathrm{x}\end{array} \right\rangle \;\; \begin{array}{|c|}\hline \mathrm{Scope} \\ \hline\end{array}}$$

With Scope as in (194).

It is now possible to justify the set of presuppositions of the preliminary representation for the first sentence of (181) as follows. We globally accommodate the assumption that the predicate C′ is one which is only true of persons in possession of a car, for instance by identifying C′ with the predicate "there is a car such that — owns". This allows us to derive from the updated global context the quantificational statement given in (196).

(196)



(196) enables us to enrich the antecedent of (195), which verifies all that the antecedent of (196) claims of the quantified variable w holds for its quantified variable x, with the information contained in the nuclear scope of (196). If we now resolve the anaphoric discourse referent y by identification with x, and the contextual predicate C by identification with "$\in$ {x,z}", then the existence-and-uniqueness presupposition in 194 is satisfied too.

The moral of this story is that even in cases like this global accommodation can produce the desired effect as well as non-global justification. We want to stress in this connection that the assumption of the contextual restriction on the quantifier expressed by *every Angelino* is independently motivated. The reason we did not display such contextual dependencies of quantifiers before is that up to now they played no part in our considerations.

The possibility of obtaining the intuitively plausible readings of sentences like (181) as the result of global accommodation is important to us, since we see the notion of non-global accommodation as conceptually problematic. When the context available to the recipient of an utterance U is insufficient for justification of all the presuppositions it generates, it is natural for him to take himself to be underinformed about the context K_S that the speaker is actually assuming (or "presupposing", in the sense of those who see presupposition as in the first instance a pragmatic phenomenon[60]) in producing U. If K_S weren't capable of justifying all presuppositons of U, then the speaker surely wouldn't have expressed himself in the way he did. On the basis of this "speaker knows best" principle the interpreter will, within a certain range delimited by further constraints on accommodation, assume that K_S is a context in which the information needed for justification is included. And if this is the rationale behind accommodation, then accommodation is an essentially global phenomenon.

[60]See [Stalnaker, 1979; Stalnaker, 1972; Stalnaker, 1974]

We end this Section with a succinct statement of the two complementary theses on presupposition justification and accommodation to which we have committed ourselves here and in the preceding Section:

(197) (General theses concerning the justification and accommodation of presuppositions in logically embedded positions)

(i) A presupposition K must be justified in its local context K_l. (But the justification may use everything in the total information of K_l, TILC(K_l,K,K_g).)

(ii) Accommodation for the sake of presupposition justification is always accommodation of the global context.

4.5 Construction of Preliminary Representations

Perhaps the greatest challenge for a DRT-based account of presupposition — as for DTR-based accounts of almost any aspect of natural language — is to formulate the rules according to which semantic representation are constructed. In the case of presupposition this challenge concerns in the first place the construction of the preliminary representations in which presuppositons are explicitly represented.

In view of the importance that representation construction has for any application of DRT, it may be felt as something of a let-down that this is precisely the part of the presupposition theory outlined here about which we will say next to nothing. Our excuse is that in order to do a proper job on this part of the theory we would have to go into much technical detail, which would detract from the more fundamental points where the present account differs from others. Also, it would have taken up so much space that little would have been left for other aspects of the theory. Given that within the present survey presupposition is only one of a substantial number of topics, the space we are devoting to it may already seem out of proportion as it is.

All we will do in this section is to outline the major issues with which a construction algorithm for preliminary representations has to cope. For further details we refer to [Kamp, 2001a] and to [Hans and Reyle,].

The first point is this. Rather than building representations from the sentences of a discourse by traversing their syntactic trees from the top down (as was done in the original formulation of DRT as, e.g., in [Kamp and Reyle, 1993]; for discussion see Section 2), we use a bottom-up algorithm.[61] It is a familiar fact from other bottom up, "compositional", definitions of sentence meaning (cf. e.g. [Cooper, 1975]), that these are often forced

[61]There are several proposals for bottom up construction of DRSs in the DRT-literature. See for instance [Asher, 1993], [Zeevat, 1989].

to make use of variable stores. (The need for variable stores arises whenever variables are introduced at one stage in the construction and bound at some later stage, with other stages in between.) This applies also to bottom up construction algorithms for DRT like the one that is at issue here. There the need for a store arises among other things for the location times of eventuality variables introduced by verbs.[62] According to the usual assumptions about syntactic structure these may get bound at a much later stage, when the construction process reaches the information contributed by tense. Many syntactic theories assume this information to be located at some functional projection of the verb fairly high up in the tree (such as Infl in pre-Minimalist versions of Chomskyan syntax), which can be at a considerable distance from the node of the verb itself. Variable storage, moreover, is also indispensible within the set up of U(nderspecified) DRT.[63]

The algorithm for constructing preliminary representations uses variable storage widely. In particular, it assumes that the discourse referent representing its referent (or, in the case of quantificational NPs, the discourse referent which plays the role of the variable bound by the quantifier) gets introduced by the head noun, but may be bound only later on and thus must be kept in store until then. "Binding" of the discourse referent introduced by the lexical head of an NP can take various forms. Binding can be quanficational, in which case the element responsible for it is the determiner of the NP; it can be effected by some other, NP-external operator, as we find with indefinites, according to the proposals of FCS, classical DRT and other forms of Dynamic Semantics; or it can take one of the various forms of referent identification that are associated with the different types of definite NPs.[64]

Among the different modes of referent identification for definite NPs there are, we have seen, in particular those which take the form of finding an anaphoric antecedent in the discourse context. Within the present discussion it is this anaphoric kind of binding that is of primary interest to us. In Section 4.1 we saw that such anaphoric binding is not only the standard form of binding for anaphoric pronouns but that it also plays a part in the interpretation of at least some definite descriptions. Moreover, it is arguable that other-than-first occurrences of proper names in a discourse involve such

[62] See e.g. [Reyle *et al.*, 2000]

[63] See [Reyle, 1993],[Eberle, 1997]

[64] In most cases, it is the determiner of the NP which tells us what kind of binding is wanted (even if it is only with quantificational NPs that the determiner then also takes care of the binding itself). With definite NPs the matter is a little more complicated, since many of these — proper names, pronouns and simple demonstratives — no clear separation between determiner and lexical head can be made. In these cases a more complicated story has to be told. For the purpose of the present discussion it suffices to assume that the single word of which such NPs are made up unites the function of lexical head (and thus variable introducer) and determiner (and thus indicator of binding mode) in one.

"antecedent" binding as well; and that the same is true for certain simple and complex demonstatives. In all such cases the algorithm under discussion stores the discourse referent, x, that is introduced by the "anaphoric" NP initially, together with information about the way in which it is to be bound when the time for binding will have come – information which depends at least in part on what NP type (pronoun, definite description, proper name, demonstrative, ...) x belongs to. The account of "antecedent" binding presented here entails that at some point the store entry for such a discourse referent must be converted into the representation of the sort of anaphoric presupposition we have encountered in the preceding sections, and this representation adjoined to that part of the representation under construction which contains the store of the given entry.

To give an impression of how the discourse referents introduced by (the heads of) anaphoric NPs are processed by the construction algorithm for preliminary representations we present a selection of the successive stages in the construction of the preliminary representation of (158.b) of Section 4.2. This will also reveal some other aspects of representation construction by this algorithm. We will not explain all details of the construction, nor of the notation used to record its various intermediate stages. The intereseted reader will have to consult the papers mentioned at the beginning of this Section.

(158.b) Every friend of mine who has a rabbit overfeeds it.

The NP *a rabbit* leads to the representation in (198).

(198) $\left\langle \left\{ \langle\ \text{y},\ \boxed{\text{rabbit(y)}},\ \text{indef.art}\ \rangle \right\},\ \boxed{} \right\rangle$

This representation consists of a variable store with one entry (for the variable y) and an empty DRS. (This DRS is to be thought of as representing the predication which involves the NP as argument. It will get filled when the representation of the NP is combined with that of its predicate — here the verb *have*. (At that point the empty DRS of (198) gets merged with that which represents the predicate and the resulting DRS is empty no longer.) The entry for u consists, as do all store entries, of three components, (i) the variable itself; (ii) a simple or complex predication of this variable, presented in the form of a DRS (also often empty); and (iii) a "Binding Constraint", which can be inferred from the source introducing the variable or its syntactic environment — here the indefinite determiner *a*. These Binding Constraints are presented here only schematically, by expressions like "indef.art". These expressions should be seen as abbreviations of the often complex binding information that a full and explicit presentation of the construction algorithm must spell out in detail.[65]

What has just been said about *a rabbit* applies mutatis mutandis also to the two other NPs, the complex Subject NP of the sentence beginning with *every*, and the direct object pronoun *it*. The representation for *it* is given in (199), that for the subject NP in (200).

$$(199)\ \left\langle \left\{ \langle \text{u}, \boxed{\text{non-pers(u)}}, \text{an.pron} \rangle \right\}, \boxed{} \right\rangle$$

$$(200)\ \left\langle \left\{ \langle \text{x}, \boxed{\begin{array}{c} \text{fr.o.m.(x)} \\ \left\langle \left\{ \langle \text{y}, \boxed{\text{rabbit(y)}}, \text{indef.art} \rangle \right\}, \boxed{\text{have(x,y)}} \right\rangle \end{array}}, \text{every} \rangle \right\}, \boxed{} \right\rangle$$

(199) is much like (198), the only difference being that its Binding Constraints are now the presuppositional ones of anaphoric pronouns rather than the indefinite Binding Constraints of *a*-NPs. The story of (200) is more complicated. To obtain this representation several construction operations are needed. Some of these are required for the construction of the representation of the relative clause, and one for the combination of that representation with that for the lexical head *friend of mine* (which for presentational purposes we treat here as if it were a single noun). The main point here is that the integration of RC representation and head noun yields a complex representation for the second component of the store entry for the subject, one which once again has the form of a DRS preceded by a variable store.

Combining the representation of the direct object with that of the verb *overfeed* yields (201) and combining that with (200) the representation in (202).

$$(201)\ \left\langle \left\{ \langle \text{u}, \boxed{\text{non-pers(u)}}, \text{an.pron} \rangle \right\}, \boxed{\text{overfeed(ARG}_1\text{,u)}} \right\rangle$$

$$(202)\ \left\langle \left\{ \begin{array}{l} \langle \text{u}, \boxed{\text{non-pers(u)}}, \text{an.pron} \rangle \\ \langle \text{x}, \boxed{\begin{array}{c} \text{fr.o.m.(x)} \\ \text{K} \end{array}}, \text{every} \rangle \end{array} \right\}, \boxed{\text{overfeed(ARG}_1\text{,u)}} \right\rangle$$

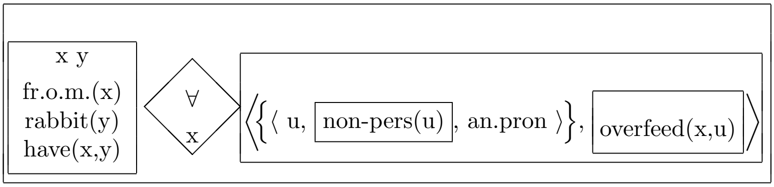
$$\text{with K} = \left\langle \left\{ \langle \text{y}, \boxed{\text{rabbit(y)}}, \text{indef.art} \rangle \right\}, \boxed{\text{have(x,y)}} \right\rangle$$

(202) can now be converted into the desired preliminary representation by *implementing* the Binding Constraints. We assume that the variable y for

[65] The spelling out of "indef.art" and other Binding Conditions is arguably the most demanding part of the entire algorithm specification. In fact, much of the linguistic literature on the semantics and pragmatics of different types of noun phrases can be seen as relevant to the exact form in which the Binding Constraints should be stated.

the indefinite gets bound locally, in the familiar DRT-mode of insertion into the local DRS universe. (The effect of this assumption is that the indefinite is interpreted as having narrow scope with respect to the universal quantifier expressed by *every*. The binding of y has the effect that the predicate occupying the second slot of its store entry gets added to the DRS whose universe receives y. In the present case this is the DRS for the second slot of the store entry for the subject NP, which becomes the restrictor of the duplex condition that results from implementing "every". These two conversions of Binding Constraints into actual bindings yield (203).

(203)



Implementation of the presuppositional Binding Constraint then yields the preliminary representation (161) of 4.2 (if we abstract of the present treatment of *friend of mine* as an atomic 1-place predicate. From this representation one then derives, by the "local" presupposition resolution described in Section 4.2, the final representation (162).

The computation of the representations of pronoun presuppositions is simple in that it has to deal with a fixed (and very limited) amount of descriptive information. With other kinds of presuppositions – including in essence all the presuppositions that are considered in the long tradition of non-anaphoric approaches to presupposition, from Frege to Heim — this is not so: For all such "traditional" presupposition types there is no upper bound to the complexity that this descriptive information can have. This is plain for factive presuppositions — the complement of a factive verb can be as complicated a sentence as you like. But it is equally true for the existence-and-uniqueness presuppositions of definite descriptions – since there is no upper bound to the complexity of the relative clauses that definite descriptions can contain — or for *again*-presuppositions, since *again* may have scope over a VP which includes NP arguments, PP adjuncts, or subordinate clauses, and for each of these categories complexity has no upper bound. Similar considerations apply to all other presuppositions which in earlier theories were treated as "presupposed propositions".

The problem that all these presuppositions present for the construction of preliminary representations was mentioned in Section 4.3: The representation of the presupposition must be obtained as a "copy" of the representation of the sentence part to which the trigger applies. (It is useful in this connection to think of the presupposition trigger as an operator whose operand is the part whose representation must be "copied" to get the representation of the presupposition it triggers. It is immaterial in this connection whether the part in question is a complement of the trigger, as with the typical factive verb, or the trigger an adjunct to the part, as we find with adverbs such as *again* or *too*. We will illustrate a couple of aspects of the copying problem for the case of *again*-presuppositions, starting with example (167) of Section 4.3.

(167) John made a mistake again.

One preliminary representation for (167) was given in (170), also repeated here.

$$(170)\quad \boxed{\begin{array}{c} \text{j} \\ \text{John(j)} \\ \left\langle \left\{ \boxed{\begin{array}{c} \underline{\text{t}} \\ \text{t} \prec \text{n} \end{array}} \right\}, \left\langle \left\{ \boxed{\begin{array}{c} \text{t}'\ \text{e}'\ \text{y}' \\ \text{mistake(y}') \\ \text{e}'\text{:make(j,y}') \\ \text{e}' \subseteq \text{t}' \\ \text{t}' \prec \text{t} \end{array}} \right\}, \boxed{\begin{array}{c} \text{e y} \\ \text{mistake(y)} \\ \text{e:make(j,y)} \\ \text{e} \subseteq \text{t} \end{array}} \right\rangle \right\rangle \end{array}}$$

To get a better grip on what is involved in the construction of such a representation we give the representation of that part of the syntactic analysis of the sentence which immediately precedes the construction stage just before the trigger *again* comes into play. We assume that *again* is an adverbial adjunct to the VP, so the representation in question is that of the VP. This representation is given in (204).

$$(204)\quad \left\langle \left\{ \begin{array}{l} \langle \text{t}, \boxed{}, \text{m.ev.l.t.} \rangle \\ \langle \text{e}, \boxed{\text{e} \subseteq \text{t}}, \text{m.ev.} \rangle \\ \langle \text{y}, \boxed{\text{mist(y)}}, \text{ia.} \rangle \end{array} \right\}, \boxed{\text{e:make(ARG}_1\text{,y}} \right\rangle$$

(204) has a store with three entries, one for the variable introduced by the direct object, and two, e and t, connected with the eventuality described by the verb, the eventuality e itself and its location time t. The respective Binding Constraints "m.ev.l.t." and "m.ev." contain information pertinent to the binding of these variables. "m.ev.l.t." — "m.ev.l.t." is short for "main eventuality location time" — abbreviates a complex set of conditions which articulate the various ways in which such location times can be bound.[66] All we need to know in connection with the present example

[66] The matter is as complex as it is, because binding of location times can take many different forms. One possibility is the indexical binding by finite tense (as we assumed in our treatment of (167) in 4.3, via the conditrion "t ⟨ n" and an additional requirement of antecedent-binding in context). But there are many other possibilities as well. Location

is that the indexically constrained anaphoric binding represented in (170) is among the options "m.ev.l.t." provides for. By comparison the Binding Constraint "m.ev." for e is simpler. We assume that e gets bound by insertion into an appropriate DRS-universe.[67]

It is from the structure in (204) that the presupposition triggered by *again* must be constructed. One question which the formulation of this operation must adress is which elements of the store require duplication and which do not. Our earlier treatment of this example took it for granted that all store elements are to be duplicated, but we will see that this is questionable. Another question concerns the eventual scope which these store elements acquire when Binding finally takes place. As we saw in Section 4.3, a precondition of the presupposition construal represented in (170) was that the variable t be bound anaphorically. This decision amounts to a kind of "disambiguation" of "m.ev.l.t", which we abbreviate "m.ev.l.t;an.". This means that t is not only related by tense to the utterance time n, but that moreover it is identified with some (past) time t_0 provided by the context. The main point of this "disambiguation" of "m.ev.l.t" there was, it enabled us to enter the temporal precedence condition "$t' \prec t$" into the representation of the presupposition, rather than into the non-presuppositional part. Formally, however, this possibility depended on the t-presupposition having wider scope than the *again*-presupposition. By ...... we would now want the *again*-presupposition to be within the scope of the store element for t.

What should we assume to be the scope relation between the *again*-presupposition and the other store elements of 204? For the present example it won't matter which way we decide. But a general principle is needed on the basis of which decisions are to be made. At this point we do not feel able to state such a principle, but even only put forward a few hints about the form it should take.

First, a correlation between the scope question and onother one which is even more important. (It matters in almost all cases, the example before us among them.) This is the question which of the elements in the store of the represenations that is in the scope of the trigger at the point when the representation is constructed for the triggered presupposition need be "copied" — i.e. whether a store element of the same form but involving a

---

times can be bound, either internally to the clause containing the verb responsible for its introduction or externally to it, via the binding relations that often exist between finite subordinate clauses and the clauses to which they are adjoined, gerundival and other infinitival constructions (including control), adverbial quantification and aspect operators, and possibly others as well. [Reyle *et al.*, 2000] give an impression of this complexity, even when restricted to possibilites of clause-internal binding. For discussions of clause-external binding see e.g. [Ogihara, 1999; Abusch, 2004]).

[67]The default assumption is that e gets inserted into the universe of the DRS which contains the condition "e:..." as one of its conditions. But sometimes there are other possibilities as well. In this regard eventuality discourse referents are much like those introduced by indefinite NPs, although we don't know how close the similarities are.

different discourse referent should be included in the store of the representation for the triggered presupposition. The natural correspondance seems to be this: Precisely those store elements of the argument representation of the representation in the scope of the presupposition trigger should be copied into the store of the representation of the triggered presupposition which remain within the scope of the new presupposition representation in the representation which results from its construction.

This correlation doesn't tell us, however, how either decision – which store elements remain within the scope of the new presupposition, which store elements must be copied — is to be made. This is a hard problem since so many different factors seem to impinge on its solution. And as with other questions of semantic scope, there appears to be room for genuine underspecification by syntactic form. The best way to deal with this and other scope problems is therefore, we believe, withing the setting of UDRT [Reyle, 1993]. Among the various constraints on the solution to the present scope problem there is one which deserves to be mentioned here, as it concerns the interaction between presuppostions. In general some of the store elements that may occur within the scope of a presupposition trigger like *again* may be persuppositional themselves. These are the store elements introduced by definite NPs which will have to be converted into presuppostion representations at some stage. (In the discussion above it was assumed implicitely that this happens at the point when the input tree to the construction process has been entirely transformed into a representation form of the kind illustrated in the representations (198)-(204) above. But for the present point it doesn't really matter when we take these conversions to take place.) What can be said about the scope relations between different presuppositions? In many cases, including all those where the presuppositions in question are resolved globally, their scope relations within the preliminary representation are of no consequence. But there are also cases where this matters. One case is discussed at length in [Kamp, 2001a]. A sentence like

(205) Fred has pawned his watch again.

is ambiguous between an interpretation according to which there was a simple watch which he pawned, then retrieved from the pawnshop and then pawned again, and a reading on which he pawned one watch, then go another one and then pawned that one too. The second interpretation can be obtained (within the present theory) only by copying the presuppositional store element which *again* finds in its scope. Rendering the possessival relation conveyed by *his* dependent on time, so that we can evaluate this relation to the time t of the asserted event in the presupposition representation adjoined to the assertion and to the time t' of the event presupposed by *again* in the copy of that representation then makes it possible to obtain two distinct referents, each of which was the unique satisfier of the relevant conditions at the relevant time. The interpretation according to which the same watch was pawned twice can also be obtained in this way, viz. when the unique satisfier of the given condition at t' is in fact the same as the unique satisfier a t. (According to this analysis the difference between the two cases isn't really a matter of two different *readings* but of two different situations to which the same semantic representation is true, which seems to be in accordance with the intuitions which some speakers have expressed about such examples like (205).

But does (205) also have another reading, which we obtain by giving the presupposition for the definite description wide scope over the *again*-presupposition? The matter is difficult to decide, since there is in principle always the possibility of making the conditions of the original presuppostion representation and its copy identical, so that they will resolve to the same referent. It is important, however, to distinguish in the present connection between definite descriptions and pronouns. In a discourse like (206)

(206) Fred hasn't got his watch on him. In fact, he has pawned it again.

the strongly preferred interpretation seems to be that the same watch was pawned twice. Anaphoric pronouns, it would thus seem, — and the same thing may well be true of anaphoric presuppositions (in the sense of 4.3) in general — come with the requirement that they "may be resolved only once". There are various ways to make sure of this within the present framework. One is to insist that anaphoric presupposition representations (or the store elements that are destined to become anaphoric presupposition representations) always are given scope over the presuppositions generated by *again* when the presupposition occurs within the trigger's scope. (Though other stipulations to the same effect are possible too.)[68]

We summarise this inconclusive discussion by stating once more the problem that it addressed: When the representation must be constructed for a presupposition-triggering particle like *again* and the representation in its scope has a store S, then the question arises which of the elements of S should get scope over the new presupposition representation and which should remain within the scope of the new representation. We assume that in general the new representation cleaves S into two parts, of the elements with wider and the elements with narrower scope. But the principles which govern this division require further investigation.

[68]The distinciton between pronouns and definite descriptions is more complex that the above remarks imply.

(207) Fred is without a watch. He has pawned it again.

can be said perfectly well in a case where Fred pawned two different watches at two different times. Examples like this one seem to be .... to the famous paycheck examples (*The man who gave his paycheck to his wife was wiser than the man who gave it to his mistress).* But exactly how this connection should be accounted for is left as a question of further investigation.

In the case of (204) we decided, in keeping with our earlier analysis of the example, that only the t-presupposition should get wide scope over the *again*-presupposition, while the other two elements of the store remain within the scope of the new presupposition, and at the same time yield copies within it. The result is given in (208).

$$(208)\ \left\langle \{\langle \mathrm{t}, \boxed{}, \text{m.ev.l.t.;an.}\rangle\}, \left\langle \mathrm{K}, \left\langle \left\{ \begin{array}{l} \langle \mathrm{e}, \boxed{\mathrm{e} \subseteq \mathrm{t}}, \text{m.ev.} \ \rangle \\ \langle \mathrm{y}, \boxed{\mathrm{mist(y)}}, \text{ia.} \ \rangle \end{array} \right\}, \boxed{\mathrm{make(ARG_1,y)}} \right\rangle \right\rangle \right\rangle$$

$$\text{with K} = \left\langle \left\{ \begin{array}{l} \langle \mathrm{t}', \boxed{\mathrm{t}' \prec \mathrm{t}}, \text{m.ev.}\mathit{ag}\text{-pr.l.t.}\rangle \\ \langle \mathrm{e}', \boxed{\mathrm{e}' \subseteq \mathrm{t}'}, \text{m.ev.}\mathit{ag}\text{-pr.} \ \rangle \\ \langle \mathrm{y}', \boxed{\mathrm{mist(y')}}, \text{ia.} \ \rangle \end{array} \right\}, \boxed{\mathrm{make(ARG_1,y'}} \right\rangle$$

"m.ev.*ag*-pr." and "m.ev.*ag*-pr.l.t." stand for "main eventuality of an *again*-presupposition" and "location time of the main eventuality of an *again*-presupposition". The Binding Constraints "m.ev.*ag*-pr." and "m.ev.*ag*-pr.l.t." are short for the special Binding Constraints appropriate for such variables.

There is a difficulty here which we passed over in our discussion of *again*-presuppositions in Section 4.2: Should we see the presupposed variable e′ and its location time t′ as existentially quantified within the *again*-presupposition, so that this presupposition has the status of a presupposed proposition? Or should one or both of them be treated as anaphoric discourse referents? Our discussion in Section 4.4 of the justification of the *again*-presupposition of Kripke's example (189.b) and its variant (190) might seem to suggest the second view. After all, in the two cases of justification that we considered in Section 4.4 the context did provide an event with which e′ could be identified (as well as a time for the identification of t′). On the basis of other cases, however, it appears to us that for the justification of an *again*-presupposition no explicit representation of an eventuality and/or location time in the context is required; it is enough if the context can be seen to entail that there was an earlier occurrence of an eventuality of the desired type. Hence no underlining of e′ and t′ in (170).[69]

4.6 Conclusion

This very brief Section serves both as a conclusion to the Section 4.5 and as conclusion to Section 4 as a whole. We extract what we see as the most salient features of the presuppositon theory presented here.

[69]It should be clear that the absence of underlining for these variables is a reflection of the Binding Constraints "m.ev.*ag*-pr." and "m.ev.*ag*-pr.l.t.". Generally, presence or absence of underlining is something that the Binding Constraints for the variables of presupposition-triggering NPs must make explicit.

1. The general approach towards the theory of presuppositions of which the first explicit version in print is [van der Sandt, 1992] and of which the present proposal is an instance, implies a sharp separation between:

(i) the *computation* of presuppositions, which is part of the construction of the preliminary sentence representations in which presuppositions are explicitly represented, and

(ii) their *justification*, which is part of the integration of the preliminary representation with the context.

This separation "presupposes" a two level DRT architecture, in which sentences are first assigned a preliminary representation which is then subsequently connected with the context representation.

In recent years it has been principally the second problem, that of presupposition justification, on which most of the work in presupposition theory was focussed. The problem of presupposition computation has often been bypassed, partly, we suspect, because systematic proposals for a syntax-semantics interface which includes presuppositional phenomena were lacking altogether. But the problem of presupposition computation should not be underestimated. There are various reasons why it shouldn't be. A particularly important one is that so often, and in the plainest and seemingly most innocent uses of language, a single sentence will give rise to several presuppositions at once.

2. Following [van der Sandt, 1992], the theory is set up to deal with phenomena which have been traditionally classified as cases of presupposition and those that have been classified as cases of anaphora in largely parallel ways. Nevertheless differences between these two kinds of phenomena remain. The present theory endeavours to do justice to these differences by distinguishing between anaphoric presuppositions and non-anaphoric (or "purely propositional") presuppositions. Whether or what further distinctions will prove necessary is a matter which we have left open.

3. The theory makes a sharp distinction between presupposition justifications that are accommodation-free (cases which in [van der Sandt, 1992] and elsewhere are described as "presupposition binding"), and cases where accommodation plays a role. One difference between justification and accommodation on which the present theory insists is that justification is always "local" and accommodation always "global".

4. The theory is designed to deal not only with single presuppositions individually but also with the (extremely common) cases where a single sentence generates several presuppositions at once and where these interact in often intricate ways. This is a domain in which there is need for much further work, both with regard to presupposition computation and to presupposition justification.

5 PROPOSITIONAL ATTITUDES

5.1 Introduction

There is a natural connection between DRT and the description of propositional attitudes, such as belief, desire or intention. The most direct connection is with belief. According to DRT, interpretation of an assertion one hears or reads takes the form of constructing a DRS for it. One way to think of this DRS is as a structure which the interpreter forms in his mind and which for him identifies the content of the interpreted statement.

In most presentations of DRT this connection is played down: as a theory of semantics, it was felt, DRT should be able to stand its ground without reference to the minds of language users. Emphasising the psychological angle would only have detracted from those aspects of the theory which make it useful as a tool for linguistic analysis in which the mental plays no direct part. The conviction that linguistics should stay clear from assumptions about what goes on in the heads of speakers or hearers was particularly strong within the context in which DRT was first developed (that of the formal semantics community of roughly twenty five years ago), and there was a correspondingly strong reluctance to dwell on the psychological potential of the theory. In the meantime, suspicion of reference to the mental has lessened even among formal semanticists. But even today it seems good policy to keep those aspects of DRT that make it a "mind-neutral" theory of meaning separate from what the theory might have to say about mental representation. This is the policy that we ourselves have followed in earlier work on DRT and to which we have also stuck in the present overview.

It should nevertheless be admitted that the idea of a mental representation which the interpreter of a sentence, text or bit of spoken discourse builds was an essential motive for developing DRT, even if the standard formulations of DRT that have made it into print bear little evidence of this. Witness to this are publications which explicitly explore the possibilities of DRT as a theory of propositional attitudes. Some of this work goes back to the eighties (cf. [Asher, 1986; Asher and H.Kamp, 1989; Kamp, 1990; Asher, 1993]).

The reason why the psychological significance of DRSs seemed a promising line of investigation from the start is directly connected with what DRT has to say about the semantics of indefinite expressions and anaphora to indefinite antecedents (highlighted by donkey sentences and donkey discourses; see Section 2.3). Suppose that a recipient B has just interpreted a sentence containing an indefinite NP α and that the next sentence he must interpret contains a pronoun for which α is a fitting antecedent. According to DRT the anaphoric connection between pronoun and NP can be established by identifying the discourse referent for the pronoun with the one for α. It is tempting to think that this account of what goes on in establishing indefinite-pronoun links tells us something about how the content of interpreted sentences is represented in the interpreter's mind: the indefinite α does give rise, at the level of mental representation, to the introduction of an entity representation (corresponding to the discourse referent for α) and this representation can then serve, just as could in principle any other entity representation in the mind of the interpreter, as an antecedent for anaphoric noun phrases occurring in sentences that are to be interpreted subsequently.

The fact that cross-sentential anaphora works in the way the theory predicts (with some exceptions, but on the whole the number of these does not seem damning), and that the theory gives such an apparently simple account for it, was one reason for thinking that DRSs capture some genuine aspect of the way in which the mind represents menal content. A further early reason for thinking this was the observation, due to Partee, that pronominal anaphora is sensitive to the form of the preceding sentence, and not just to its "propositional" (i.e. intensional) content: "It is under the sofa." can be understood as a statement about the missing marble when it follows "One of the ten marbles is not in the bag." but not when it follows the propositionally equivalent "Nine of the ten marbles are in the bag". (See section 3.1 This distinction is also captured effortlessly by DRT, and it is one which seems to go directly against the fundamental assumptions about semantic content that were dominant within formal semantics at the time.

Even if these and other facts (some discovered by psycholinguists over the past twenty years) make it plausible that entity representation (including representation of entities introduced by indefinite NPs) has psychological reality, we must, when it comes to claiming psychological reality for the representational form of DRSs generally, tread very carefully. About the mental representation of predication very little is apparently known even today. Thus it would be premature to consider all aspects of the form of DRSs as capturing aspects of psychological reality.

In this section we will discuss an extension of DRT, in which DRSs will be used to identify mental representations of content. We want to remain agnostic, however, on the question precisely which features of DRSs are psychologically significant and which are not, leaving these questions to be settled by future work in cognitive science.[70] We certainly do not advocate

[70]In this respect the account presented here is less committal than, for instance, [Asher, 1986], where the form of DRSs is used to arrive at an account of the identity conditions of beliefs and other propositional attitudes. Our own inclination on this point is that the concept of identity for beliefs and other propositional attitudes is too context-dependent to allow for a characterisation once and for all in any case.

wholesale adoption of the DRS-format as psychologically significant in each and every respect.

5.2 Extending DRT to a Formalism Capable of Describing Attitudinal States and Attitude Attributions. Some Examples Semi-Formally Treated

As indicated above, a principal motive for applying DRT to the analysis of mental contents is its ability to deal with cases of cross-sentential donkey anaphora and the way in which it does this: the new sentence with the anaphoric pronoun is interpreted via a representation in which the discourse referent of the pronoun is identified with that of its antecedent. A consequence of this is that the DRS K_2 for the new sentence is not a proper DRS; one of the discourse referents occurring in conditions of K_2, viz. the discourse referent for the pronoun's antecedent, is not bound within K_2 itself, but in the DRS K_1 which represents the preceding sentence or sentences and serves as context of interpretation for the new sentence.[71] In standard DRT the non-properness of K_2 does not cause problems, since what counts in the end is only the merge of K_1 and K_2, and that DRS will normally be proper even if K_2 isn't.

For the question how content is mentally represented, cases of cross-sentential anaphora to indefinite antecedents hold a double moral. First, if the representation of content is along the lines DRT describes, then representation of new information, and thus of the content of newly acquired propositional attitudes, will take the form of "pegging" the new representation on one that is already in place. By itself the new representation would not determine a well-defined propositional content; it succeeds in doing so only in conjunction with the representation of some other attitude, on which it depends "referentially". Let us assume that the recipient of a two-sentence discourse in which the second sentence is in such an anaphoric dpendence on the first sentence, that both sentences are asertions which communicate new information to him and that the recipient accepts both bits of information as true and thus forms the corresponbding beliefs. It is natural in such a situation to think of the first DRS, K_1, which (we assume) the recipient has formed as the result of his interpretatiojn of the first sentence as representing for him the content of the first of his two beliefs, and of the sewcond DRS, K_2, the result of his interpreting the second sentence, as representing the content of the second belief. So far the storz may seem barely distinguishable form the one whic standard DRT tells about incremental interpretation of discourse. But there is one difference with what

[71]Note that this is so irrespective of whether we insert the discourse referent α for the antecedent in the argument slots of the pronoun, or proceed as we have here, viz. by introducing a separate discourse referent β for the pronoun and then add to K_2 the condition "$\alpha = \beta$".

we have been assuming so far, and it is a difference that is crucial. For in the presnt context it is no longer possible to simply amalgamate the new, improper representation K_2 with the DRS K_1 on which it depends. The reason is that the attitude whose content is represented by K_2 may be of a different kind from the one represented by K_1. For instance, K_2 may represent a belief with a lower confidence degree than K_1. Or, more dramatically, the attitude represented by K_2 could be a desire while that represented by K_1 is a belief; and so on. To give an example of the first case, consider a sitaution in which two philosophers, Φ_1 and Φ_2, are talking over coffee. Φ_1 is telling Φ_2 about the last convention he went to, and which Φ_2 had decided not to attend. "You know", he says to Φ_2,

(209) "I gave my paper on implicature, the one you have seen. There was one person in the audience who objected — of course I was prepared for an intervention of that kind — that not every case of inference is a case of implicature. Well, I wiped the floor with him."

Let's assume that Φ_2, in accordance with the speaker's referential intentions, interprets *him* as anaphoric *to one person in the audience.* This makes her representation of the third sentence in (209) referentially dependent on her representation of the first sentence. But let us assume also that Φ_2, while seeing no reason to doubt that the first two sentences of (209) are true, doubts the truth of the third: she knows Φ_1 as a rather inflated person, who tends to be out of touch with reality where his ability to convince or impress others is concerned. So she registers the first two sentences as belief (with a high confidence degree), and the third sentence as a doubt. The two representations must be kept separate, one as specification of the content of a belief and the other as specification of the content of a doubt; amalgamating them would obliterate the crucial demarcation between what is accepted as true and what isn't. It follows from this that DRS-merge can no longer be used to account for the binding problem connected with cross-sentential anaphora.

Our second example shows that the problem illustrated by the first is not restricted to attitudes which arise through the interpretation of language.[72] A stamp collector opens the lid of a box which contains an unsorted miscellany of stamps. He has been told he can pick one stamp out of the box and keep it. He perceives (or thinks he does) a copy of the 2d stamp of the 1840 edition of Great Britain (showing the head of Queen Victoria), but of which only a tiny portion is visible to him. (Stamp collectors are known to develop an uncanny ability to identify stamps even if only a tiny part of them is exposed to view.) The 1840 2d of GB is a stamp for which he is always on the look-out. So his perceptual experience instantly produces in him (i) the belief that there is a copy of this stamp in the box, (ii) the

[72]See [Kamp, 2003].

desire to make this copy his own; and (iii) the resolve to pick the stamp out of the box (thereby making his desire true).

Let us assume that each of these attitudes can be represented as a pair consisting of (a) a representation of its propositional content and (b) an indicator of its *attitudinal mode* — that is, some feature which distinguishes between beliefs, desires, intentions, etc. For simplicity we will assume just three such mode indicators here — BEL, for belief, DES, for desire, and INT, for intention. This way of representing propositional attitudes "two-dimensionally" allows among other things for the possibility that the same propositional content representation can combine with different *mode indicators*. This corresponds to the possibility of representing attitudes with the same content but distinct modality, as when two different persons hear the same assertion and assign the same interpretation to it, but where one accepts it as true, while the other withholds judgement; or when a person first believes something but then discovers this belief to be false; or when someone has a fervent hope that something is the case (e.g. that his beloved is still alive) and then finds his hope confirmed. The representation of attitudes which coincide in content but differ in mode is especially important in the description of dialogue — where the participants will often have different attitudes towards the same proposition — and also in describing how attitudinal states change in time, either under the influence of new incoming information or as a consequence of internal information processing (i.e. reasoning).[73]

Given these assumptions, and ignoring for the moment the temporal dimension, we can represent the complex of the three attitudes that, in our presentation of the case, result from the (presumed) perception of the Queen Victoria stamp as follows:

$$
\begin{array}{ll}
 & \left\langle \text{BEL}, \boxed{\begin{array}{c} \text{x z} \\ \text{2d-1840-GB(x)} \quad \text{(the box)(z)} \quad \text{in(x,z)} \end{array}} \right\rangle \\[2ex]
(210) & \left\langle \text{DES}, \boxed{\begin{array}{l} \\ \text{poss(i,x)} \end{array}} \right\rangle \\[2ex]
 & \left\langle \text{INT}, \boxed{\begin{array}{l} \\ \text{pick-from(i,x,z)} \end{array}} \right\rangle
\end{array}
$$

The first point connected with (210) is that the discourse referent x for the stamp, which is bound in the belief-component (through its presence

[73] The possibility of representing attitudes in this way is also important for accounts of belief revision which pay closer attention to the form in which new information becomes available than is done in the classical approaches to belief revision. (For the classical approach see e.g. [Gärdenfors, 1988].)

in the universe of the belief DRS), recurs in the desire and the intention components. So the DRSs of these components do not determine a well-defined propositional content without support from the belief DRS. The second point, to which we will return below, concerns the symbol "i". "i" is specific to attitude representations and there it stands for the "self", i.e. for the bearer of the attitude as he immediately perceives himself, in particular as the subject of his own perceptions and actions. Thus "i" acts as an indexical discourse referent. A token of "i" which occurs in the representation of an attitude of an agent A will ipso facto stand for A.[74]

The third point is that the "referential sharing" between the belief, the desire and the intention of (210) which is captured by the occurrences of the same discourse referent x in each of them, is a decisive factor in the way in which someone whose mental state includes these attitudes may be expected to "act" — internally (i.e. in thought) as well as externally (i.e. in acting upon his environment). The belief that there is a 2d stamp from the 1840 edition of Great Britain in the box, we suggested, gives rise to the desire to be in possession of "that stamp", and then to the intention to take it out of the box. And the action into which this intention is likely to result — that of reaching for a stamp assumed to be the 2d. 1840 Queen Victoria and seen to be at a certain place in the box — will be guided by this intention (in combination with a further belief, or aspect of the displayed belief, which concerns the precise location of the stamp in the box, one that also comes from the perception, but which we didn't display in (210)). The desire and intention develop out of the belief as desire and intention *about the same object the belief is about.* And here the sense of "the same object" is clearly a psychological one, which controls the internal development of thought and its eventual manifestations through action.

In fact, it is important to distinguish this internal sense of "same object"

[74]In this regard "i" is reminiscent of the indexical discourse referent "n". We saw in Section 3.5 that an occurrence of "n" always stands for the utterance time of the sentence represented by the DRS containing it. In fact, as we will explain presently, "n" also has a use within the representations of propositional attitudes. However, there are also important differences between "i" and "n".

Representations of beliefs and other attitudes which contain "i" should be distinguished from representations which contain in lieu of "i" a non-indexical discourse referent x which (as it happens) represents the thinker himself. Such a "non-first person" discourse referent x can be internally and externally anchored to the thinker (for anchoring see Section 5.3), but the anchors may be such that they do not enable the thinker to realise that he himself is the individual represented by x. In other words, these anchors do not enable him to make the transition from the thought representable as "P(x)" to the one representable as "P(i)". Kaplan's well-known case of the man who sees a person in the mirror whose trousers are on fire, who doesn't at first realise that he is that person, but for whom the penny then drops, can be described in our DRS language by a sequence of two attitudinal states, the first containing a belief representable as ⟨ {y}, {trousers-of(y,x), on fire(y)} ⟩, with x internally anchored and externally anchored to the man, while the second state, which supplants the first when the penny drops, contains instead a belief represented as ⟨ {y}, {trousers-of(y,i), on fire(y)} ⟩.

from the external sense which is prominent in many philosophical discussions of meaning, reference and the content of propositional attitudes. These discussions typically focus on cases where two expressions refer to the same real world entity, but where a particular speaker (or thinker) may be unaware of this, or alternatively where a speaker takes two expressions (or occurrences thereof) to refer to the same thing although in actuality they refer to distinct real world entities.[75] The internal sense of "sameness of reference" that we are dealing with here is different. The difference comes out clearly when we consider cases of misperception. Suppose first that our stamp collector has falsely concluded that a particular stamp, of which he glimpsed a corner, was the 2d 1840 of Great Britain. In this case the belief he forms can be construed as a false belief about a particular object (i.e. the stamp whose corner he misinterpreted). The desire and intention could then also be directed towards this stamp, even though the collector would not have had these attitudes if he wasn't under this misconception. If he implements his intention by taking the stamp out of the box, he will be disappointed. But the process leading up to this action will be, from an internal, psychological perspective, just as in the first case.

Misperception, however, can also be more radical than this. The belief that there is a specimen of the 2d 1840 of Great Britain in the box may have been caused by some combination of optical factors that led to this illusion without there being any one stamp in the box that is directly responsible for it — there is no stamp of which it could be said that the observer had misidentified that stamp as the 2d 1840 of GB. In such cases it is plausible to hold that the belief — and with it the desire and the intention — are not about any one object in particular. This is the position we adopt in relation to cases of this latter sort. In fact, we will argue below that if the sitatioj is as in this last scenario, then the representations in (210) fail to define a propositional content altogether (although there are closely related representations which are also within reach of the agent and which define the corresponding existential propositions, e.g. the proposition that there is a specimen of the 2d 1840 of GB in the box).

We assume that the crucial difference between this last case and the first two is that the discourse referent x has an external anchor in the first two cases, but not in the last one. What we mean by this can be explained as follows. Our point of departure is a causal theory of perception according to which direct perception of an object involves a certain kind of causal relation between the thing perceived and the perceiver.[76] In the present context this view takes a slightly different form from the one in which it is normally presented: The causal relation is a three-place relation, involving (a) the object perceived, (b) the perceiver, and (c) the discourse referent which

[75] See, as one among many places in the philosophical literature, [Kripke, 1979].
[76] [Grice, 1961], [Kaplan, 1989]

arises as a constituent of the content representation of the propositional attitude to which the perception gives rise and which represents the object in that representation. So, in the first two cases considered, the terms of the relation will be: (a) the stamp whose corner suggests that it is a specimen of the 2d 1840 of GB, (b) the collector, and (c) the discourse referent x shown in (210). Since the discourse referents at issue will always be constituents of the representations of propositional attitude contents, and the perceiver is uniquely determined as the one who has this attitude, it is legitimate to talk about the causal relation as one which involves just the perceived object and the discourse referent to which its perception gives rise; and this is the practice we will adopt.

Whether an external anchor actually exists is something for which the observer cannot have conclusive evidence — this is just what the examples of optical illusion show. In the cases we just discussed, however, the observer is persuaded that he *is* seeing a particular stamp — this is as true in the third case as it is in the first two. From his own, internal perspective the three cases look exactlz alike. In each of them he takes himself to have direct perceptual contatc with an object, and about which he then forms certain beliefs, as wellas (as in the cases before us) certain other attitudes. we consider this aspect of the resulting attitude complex — that the perceiver forms a representation of something to which he takes himself to stand in direct perceptual contact — an important feature of the nature of mental representation. We capture this feature — the presumption connected with an entity representation that it is the result of a causal interaction between the one whose representation it is and that which it is presumed to represent — in terms of the notion of an *internal anchor*. Internal anchors are, unlike the external anchors we have just spoken of, constituents of the mental state of the perceiver. We represent them as separate components of the attitudinal state as a whole, on a par with those constituents that are genuine propositional attitudes.

We assume that internal anchors carry some information about how the anchored discourse referent anchored. Thus a perceptual anchor like the one for x in our example will contain some information which records how the perceiver perceives (or thinks he perceives) the represented object. We will not be very precise about exactly what information should go into internal anchors, leaving this as a question for further research.It should be stressed, however, that we regard perception as only one of several sources of anchoring. Other causal relations between a cognitive agent and an object can also give rise to anchored representations. And in these cases too the anchored representation may be legitimate (i.e. it did arise from an actual interpretation between the agent and the represented object) or it maz be the illegitimate product of an illusory interaction, in which case there will be, once again, an internal anchor but no external one. The anchoring information that is part of such non-percpetual anchors will of course be different that which is part of the various kinds of perceptual anchors. But, as said, this is an aspect of the conecpt of an anchor about which we remain neutral. As an indication that we admit other anchors besides the perceptual ones, we will occasionally mark the latter with the subscript "dir. perc.".

With regard to our example of the stamp box we assume that (in all three scenarios) the agent has anchored representations not only for the (presumed) stamp, but also for the box.[77]

Thus we now get a total of five components instead of the three of (210).

(211)

$$\left\langle \text{[ANCH,z]}, \begin{array}{|c|}\hline \text{z} \\ \text{(the box)(z)} \quad \text{infrontof(z,i)} \\ \hline \end{array}_{\text{dir.perc}} \right\rangle$$

$$\left\langle \text{[ANCH,x]}, \begin{array}{|c|}\hline \text{x} \\ \text{stamp(x)} \quad \text{in(x,z)} \\ \hline \end{array}_{\text{dir.perc}} \right\rangle$$

$$\left\langle \text{BEL}, \begin{array}{|c|}\hline \\ \text{2d-1840-GB(x)} \\ \hline \end{array} \right\rangle$$

$$\left\langle \text{DES}, \begin{array}{|c|}\hline \\ \text{poss(i,x)} \\ \hline \end{array} \right\rangle$$

$$\left\langle \text{INT}, \begin{array}{|c|}\hline \\ \text{pick-from(i,x,z)} \\ \hline \end{array} \right\rangle$$

The representation in (211) leaves open whether the internally anchored discourse referents x and z are also *externally* anchored. This is information that, as noted above, cannot be part of an attitude description of which

[77] The anchored discourse referent of an internal anchor is mentioned not only in the universe of the DRS that occurs as second component of the anchor, but also in the first component. This is in order to make explicit that it is this discourse referent (i.e. the one which occurs as part of the first component of the anchor) for which the anchor is an internal anchor. The DRS universe might contain additional discourse referents needed to express the anchoring information which the DRS serves to represent. In that case confusion maz arise as to which discourse referent is actually being anchored. This is a complication which doesn't affect any of the examples which will be considered in this section. But it is not too difficult to come up with cases in which it does (given plausible assumptions about the information that goes into the second part of an internal anchor).

all constituents are intended to correspond to psychologically significant aspects of the represented attitude complex. It is nevertheless possible, however, for an external observer O who attributes a certain mental stateto some agent A to judge that A did truly perceive a certain object, and that his representation x of that object therefore has not only an internal but also an external anchor. It should be possible for our formalism to represent such judgements. That is, it should be possible to represent the judgement that A's representation x is externally anchored as part of O's representation of his attribution to A. We once more use our example of the Queen Victoria stamp to show how this additional information is represented in the present framework. Suppose that (211) is O's representation of the mental representation he attributes to A. O's judgement that, e.g., the discourse referent x is externally anchored, is expressed in the form of a pair $\langle$x,s$\rangle$, where s is a designator which O uses to refer to the stamp which he assumes is the perceptual origin of A's internally anchored representation. As indicated above, this information should be kept separate from that part of the representation which "models" A's mental state insofar as it is accessible to A himself. So the pair $\langle$x,s$\rangle$ is treated as part of a distinct component, and is placed to the right of the "internal" part of the representation given in (211). (212) rerpesents an attribution by O to A whose "psychological" component is like the representation in (211) and in which both the discourse referents x and z are externally anchored. (Thus O also takes A's box representation z to be the result of a true perception, and uses the designator b to refer to the box which he takes A to have perceived.)
The set $\{\langle$ z,b$\rangle$, $\langle$x,s $\rangle\}$ is called the external anchor of (212), and its member pairs $\langle$z,b$\rangle$ and $\langle$x,s$\rangle$ external anchors for z and x, respectively. If the observer O believes z to be externally anchored, but not x, then the external anchor of his description would contain the pair $\langle$z,b$\rangle$, but no pair for x. In the unlikely event that he thought even z to be the effect of an optical illusion, the external anchor would be empty; and so on. [78]

Let us return to the purely internal representation (211). The mental state represented by (211), we said, could arise in each of the three scenarios we have described. In the first of these, where the stamp of which the collector sees a small corner is indeed a 1840 2d. of Great Britain, the discourse referent x is externally anchored and the belief involving it is true. In the second scenario we still have an external anchor for x, but the belief is now false. In the third scenario there isn't even an external anchor for x. What are we to say in this case about the belief of (211)? Is it false again? Or odes its truth value depend on whether there is a specimen of the stamp somewhere in the box, or in that part of it where the collector thought he

[78] The formalism we present here does not provide for statements which deny external anchorage for a discourse referent occurring as a constituent in the described mental state. There are no principled objections, however, to extending the formalism with such means.

(212)

$$\left\langle \text{[ANCH,z]}, \underset{\text{dir.perc}}{\boxed{\begin{array}{c} \text{z} \\ \text{(the box)(z)} \quad \text{infrontof(z,i)} \end{array}}} \right\rangle$$

$$\left\langle \text{[ANCH,x]}, \underset{\text{dir.perc}}{\boxed{\begin{array}{c} \text{x} \\ \text{stamp(x)} \quad \text{in(x,z)} \end{array}}} \right\rangle$$

$$\left\langle \text{BEL}, \boxed{\begin{array}{c} \\ \text{2d-1840-GB(x)} \end{array}} \right\rangle$$

$$\left\langle \text{DES}, \boxed{\begin{array}{c} \\ \text{poss(i,x)} \end{array}} \right\rangle$$

$$\left\langle \text{INT}, \boxed{\begin{array}{c} \\ \text{pick-from(i,x,z)} \end{array}} \right\rangle$$

$$\{\langle \text{z,b} \rangle, \langle \text{x,s} \rangle\}$$

saw such a specimen? Our position is that in this last case the represented belief is neither true nor false: Since there is no particular object to which x is directly linked, and which it could thereby be considered to represent, there is a fortiori no way to decide whether or not this putative object has a certain property. Failure of an internally anchored discourse referent to have a corresponding external anchor is a failure of presupposition, which renders the question of truth or falsity moot. representations of presuppositional attitudes which contain occurrences of discourse referents that are anchored internally but lack an external anchor, cannot be avaluated as true or false; they do not determine well-defined propositions.

This position, that attitude representations with internally but not externally anchored discourse referents do not express propositions, is connected with another one. Suppose an attitude representation contains occurrences of internally anchored discourse referents but that all those discourse referents do have corresponding external anchors too. In that case the representation does determine a well-defined proposition. But the proposition expressed is a *singular* proposition. In case the representation contains just one externally and externally anchored discourse referent, this proposition is the one which attributes to the object to which the discourse referent is externally anchoredthe property expressed by the remainder of the representation. In case there are two anbchored discourse referents , the proposition attributes to the external anchors of these discourse referents a certain binary relation, and so on for numbers greater than two. In particular, in the case of the first two stamp scenarios the belief representation in (211) expresses the proposoition that says of s that it is a specimen of the 2d 1840 of GB and the representation of the intention attributes to s, b and the perceiver himself the relation which holds between any individuals a, b and c iff c takes a out of b.

The position that external anchoring entails propositional singularity while absence of an external anchor for an internally anchored discourse referent entails failure to determine propositional content can be summarised as follows:

- If all internally anchored discourse referents that occur in the representation of a propositional attitude are externally anchored, then the representation expresses a proposition that is singular with respect to each of the external anchors for these discourse referents;

- If the representation contains an occurrence of a discourse referent that is internally but not externally anchored, then it doesn't express any proposition at all.

This position might be thought to undermine the very purpose of the proposal we are in the process of developing, viz. that different propositional attitudes can be “referentially connected” by sharing one or more discourse referents — just as this is often found with the representations of different sentences in a coherent discourse. For instance, in the example on which we have concentrated so far, we considered two possibilities for the discourse referent x: either x is externally anchored, in which case each of the three attitude content representations in (211) defines a singular proposition on its own (that is, none of them needs any of the others to determine the proposition it expresses); or else x is not externally anchored, in which case none of these representations express a proposition.

So it looks like we are left with these two possibilities: either each of the components of the representation of an attitude complex defines a (often singular) proposition on its own, in which case the referential connections between them are mediated by external referents; or the "dependent" representations don't have a proper propositional content, so there are no propositional contents to be connected. Are we to conclude that the internal referential dependencies illustrated in (211) are a red herring?

That would surely be the wrong conclusion. Internal referential connectedness is a psychologically real and important aspect of thought. To repeat once more, from the internal perspective of the perceiver-agent it makes no difference whether his internally anchored discourse referents are externally anchored or not. In either case his thoughts and actions will be the same. To return to our example: The agent A will make a move for the stamp which (he thinks) he has perceived, whether or not there really is a particular stamp that has caused his visual experience. The differences between the three cases solely concern the actual outcome of the action he performs. When there is a stamp he does perceive and this stamp has the properties he perceives it to have, things will work out as he expected; in the second case he will find to his disappointment that the stamp on which his action is targeted isn't the one he thought it was; and in the third case he may come to realise that what he thought was a specimen of a certain stamp wasn't really anything at all. But while there will be variation in the result of the action, the mental process which leads to it, as well as all or most[79] of the actual motions which the action involves, will be the same.

The conclusion can only be that an account of what people think and (try to) do must be independent of whether or not the discourse referents of their attitudes are externally anchored. What does matter is how they are internally connected. What we need, therefore, is not only a semantics for attitude descriptions which take external anchors into account, but also

[79]Of course, as soon as the action leads to the agent's discovery that the stamp is different from what he thought it was or that there was no particular stamp that he saw at all, the remainder of his action may be expected to be different from what it would have been if he had perceived a specimen of the stamp of which he thought he saw a specimen. But as soon as such a discovery is made, A's internal representation of the stamp will no longer be the same, but will be modified to reflect his new information. In particular, the discovery that there was no particular stamp at all will have the effect that A's representation for the particular stamp will be expunged.

one in which only the internal properties of the represented mental state are taken into consideration. It is this second semantics which provides the basis for a useful theory of practical reasoning, not the first. What this second semantics, in which external anchors are ignored, is like, is another matter. One of the principal concerns of this section is to find out what such a semantics could be like.[80]

Attitude Attributions

There are two important features of mental states and their ascriptions which we have not yet considered. The first is the time at which the bearer of an attitude is supposed to have it. It is a crucial fact about people (and presumably certain other creatures too) that they have propositional attitudes. It is almost equally important that they can change them. People learn, part of which is that they acquire new beliefs; they forget things, they may change their agendas, i.e. their desires and intentions, and sometimes they come to realise that certain things they believed are false (a point already made). Therefore we want to be able to describe bearers of propositional attitudes not just as (timelessly) having beliefs, desires, etc. but also as coming to believe a given proposition at a certain time t, as having lost or abandoned a belief by a certain time (whether through sheer forgetfulness or by losing conviction that it is correct); and so on.

The second feature that is still missing from our representation format concerns the integration of *ascriptions* of attitudinal states into the general representation format of DRT. For instance, what would a DRS be like which combines the following bits of information: (a) that A is a stamp collector; and (b) that A has (at some particular time) the complex of attitudes represented in (212)?

We deal with these two problems — temporal dependence and integration – by one and the same representational device. It consists in introducing a special predicate, Att, into DRT's vocabulary. Att has three arguments: (i) for the individual to whom an attitude complex is attributed, (ii) for the attitude complex that is attributed to this individual and (iii) for the external anchor for this complex. Thus Att(a, K, EA) says that a is in a mental state which contains the attitudes represented in K, and that EA externally anchors some or all of the internally anchored discourse referents occurring in K. We will assume that K is a set of attitude descriptions

[80] The distinction between these two kinds of semantics for attitude representations has been discussed extensively within the philosophy of mind. There one is often led to draw the disc tinction between *narrow content* and *wide content.* This distinction corresponds fairly closely to the semantics of attitude representations with or without external anchors: narrow content ignores external anchors whereas wide content takes them into account. The correspondence is far from perfect, however, since the distinction between *narrow content* and *wide content* glosses over the problem of referential connectedness. (See [Loar, 1988; Loar, 1987; Stalnaker, 1990].)

of the kind encountered in our examples: pairs $\langle$MOD,K$\rangle$, where MOD is a mode indicator (here: BEL, DES, INT) and K is a DRS. We refer to such sets as *Attitude Description Sets* (ADSs). As we have seen, the DRSs occurring as second members of pairs belonging to an Attitude Description Set K may contain free occurrences of discourse referents so long as these discourse referents are bound in a DRS occurring as second component of some other pair in K. External anchors are as described: pairs consisting of a discourse referent occurring in K and a discourse referent not occurring in K. (The latter discourse referent functions as a K-external representation for the external anchor of the former.)

What DRS-conditions the predicate Att enters into depends on the DRS language to which it is added. However, if Att is to serve also representing the temporal aspects of attitude attributions, then the language we need is one capable of making predication time explicit generally — what we need is the language defined in Section (3.5). In this language the time of a given predication (that it is expressed by representing the predication as an eventuality to the effect that the predication holds and then locating this eventuality in time by adding further conditions (such as, say, "t $\subseteq$ s", where s is the eventuality and t its location time).

is at or during time t that the predication holds) can be expressed by locating the state which consists in the predication holding as including t. This is the device of which we also make use here. Thus, DRS-conditions involving Att will come in the form given in (213)

(213) s:Att(a,K,EA)

The temporal dimension of such predications can now be expressed by relating s to some "location" time t, which can then be further specified in various ways (see Section (3.5)). For instance, the DRS which expresses that the individual represented as a is at the time n in a mental state which contains the belief, desire, intention and internal anchors of (211) as components takes the form given in (214).

Here EA could for instance be the external anchor of (212).[81]

It should be intuitively clear how this formalism can express more complex temporal information about attitudinal states. For instance, suppose that stamp collector A's desire to have the stamp he has spotted and his intention to pick it out of the box do not arise instantaneously, and thus not simultaneously with the belief that this stamp is a specimen of the 2d 1840 of GB, but that the belief, the desire and the intention come about at three successive times t_1, t_2 and t_3. Let K_1 consist of the first three components of (211), K_2 of the first four and K_3 of all five. And let EA be the external anchor given in (212). Then an approximation of the situation described is given by the following DRS (215)

[81] N.B. When the third argument EA of Att is empty, we will usually suppress it altogether: we write "s:Att(a,K)" instead of "s:Att(a,K,$\emptyset$)".

$$
(214)\quad
\begin{array}{|c|}
\hline
\text{t s a} \\[6pt]
\text{t} = \text{n} \quad \text{t} \subseteq \text{s} \quad \text{collector(a)} \\[6pt]
\text{s:Att(a,} \left\{
\begin{array}{c}
\left\langle \text{[ANCH,z],} \begin{array}{|c|} \hline \text{z} \\ \text{(the box)(z)} \quad \text{infrontof(z,i)} \\ \hline \end{array}_{\text{dir.perc}} \right\rangle \\[12pt]
\left\langle \text{[ANCH,x],} \begin{array}{|c|} \hline \text{x} \\ \text{stamp(x)} \quad \text{in(x,z)} \\ \hline \end{array}_{\text{dir.perc}} \right\rangle \\[12pt]
\left\langle \text{BEL,} \begin{array}{|c|} \hline \\ \text{2d-1840-GB(x)} \\ \hline \end{array} \right\rangle \\[8pt]
\left\langle \text{DES,} \begin{array}{|c|} \hline \\ \text{poss(i,x)} \\ \hline \end{array} \right\rangle \\[8pt]
\left\langle \text{INT,} \begin{array}{|c|} \hline \\ \text{pick-from(i,x,z)} \\ \hline \end{array} \right\rangle
\end{array}
\right\} \text{, EA)} \\
\hline
\end{array}
$$

$$(215)\quad \begin{array}{|c|}\hline \mathrm{t}_1\ \ \mathrm{t}_2\ \ \mathrm{t}_3\ \ \mathrm{s}_1\ \ \mathrm{s}_2\ \ \mathrm{s}_3\ \ \mathrm{a} \\ \\ \mathrm{t}_1 < \mathrm{t}_2 < \mathrm{t}_3\ \ \mathrm{t}_1 \subseteq \mathrm{s}_1\ \ t_2 \subseteq \mathrm{s}_2\ \ \mathrm{t}_3 \subseteq \mathrm{s}_3\ \ \mathrm{collector(a)} \\ \mathrm{s}_1\mathrm{:Att(\ a,}\ K_1\mathrm{,\ EA)}\ \ \mathrm{s}_2\mathrm{:Att(\ a,}\ K_2\mathrm{,\ EA)}\ \ \mathrm{s}_3\mathrm{:\ Att(\ a,}\ K_3\mathrm{,\ EA)} \\ \hline \end{array}$$

This is arguably not quite what we want, since it provides no temporal lower bound for the new attitudes. For instance, (215) doesn't exclude that A had the desire to be in possession of the stamp s already at time t_1. The information that t_2 was the first time at which A had this desire can be expressed by adding the condition that a was not in a mental state of the type of s_2 at any time preceding t_2. The addition would take the form given in (216)

$$(216)\ \neg \begin{array}{|c|}\hline \mathrm{t}\ \ \mathrm{s} \\ \\ \mathrm{t} < \mathrm{t}_2\ \ \mathrm{t} \subseteq \mathrm{s} \\ \mathrm{s:Att(\ a,}\ K_2\mathrm{,\ EA)} \\ \hline \end{array}$$

This may not be the most elegant way to express such negative information, but in the formalism presented here it is the only way. A more convenient notation could be added without difficulty, if desired.

5.3 Syntax and Semantics of the Extended Formalism

Syntax

About the syntax of the extended formalism we can be brief, since all that is important has already been said.

We take the DRS-language L of Section (3.5) as our point of departure.

DEFINITION 62. (Syntax for DRS languages capable of describing propositional attitudes and attitudinal states).

Let L be the language defined in Section (3.5) or some extension of that language.

1. The vocabulary of the language L_{PA} is the vocabulary of L **(????)** together with the following two additions:

   (i) the indexical discourse referent i;

   (ii) the predicate Att.

2. The set of DRS conditions is extended via the clause:

   If s is a state discourse referent, a a discourse referent for individuals or sets thereof, K an Attitude Description Set and EA an External Anchor Description for K. Then s:Att(a, K, EA) is a DRS condition.

The notions used in 62, that of an*Attitude Description Set* and that of an *external anchor* for such an ADS, have been described infromally above. But more precise characterisations are needed. The notion of an Attitude Description Set is based on a set MI of mode indicators. In the presentation here we have opted for the set {BEL,DES,INT}, but this restriction has no direct consequences for the definition: members of an ADS are pairs of the form $\langle$MOD,K $\rangle$ where MOD $\in$ MI and K is a DRS (of the new, extended language L_{PA}). The one type of element of ADSs that requires more attention is that where the first member of the pair is an expression of the form "[ANCH,ξ]", for some discourse referent ξ. Here too the second component of the pair is a DRS. As indicated above, the repertoire of conditions that occur in such a DRS should probably be restricted to conditions (or condition combinations) of special forms. This is a matter which we ignore here. But another aspect of the pairs $\langle$[ANCH,ξ],K$\rangle$ that can occur as members of ADSs is going to be relevant later on and needs to be stated explicitly: We assume in general that the discourse referent ξ is a member of U_K. (cf. example (212) above).

Thus we come to the first of the two supplements that are needed to turn the definition of the syntax of L into a fully explicit definition of the syntax of L_{PA}:

3. An ADS of L_{PA} is a set of pairs each of which has one of the following two forms:
   (i) $\langle$MOD,K$\rangle$, where MOD $\in$ {BEL,DES,INT} and K is a DRS of L_{PA}.
   (ii) $\langle$[ANCH,ξ],K$\rangle$, where ξ is a discourse referent and K is a DRS of L_{PA} such that $\xi \in U_K$.

What remains is the defintion of the notion of an external anchor EA for an ADS *K*. But this is easy. Each ADS *K* has a set IA(*K*) of internal anchors. (These are just the members of *K* whose first component is of the form "[ANCH,ξ]".) An external anchor for *K* is simply a function whose domain is a subset of IA(*K*):

4. Let *K* be an ADS. An external anchor for *K* is a function f such that Dom(f) $\subseteq$ IA(*K*)(= {*x: for some DRS K,* $\langle$*[ANCH,x],K*$\rangle$ $\in$ K.

Semantics

The semantics for languages like L_{PA} presents us with a real quandary. The problem is a very fundamental one, and it is one which doesn't have anything to do with Dynamic Semantics as such, although one might hope that a representational approach towards semantics like that of DRT would help to find a solution for it.

The problem can be apostrophied as the gap between intentionality and intensionality. As discussed in Section (3.2), Intensional Semantics is that approach to the theory of meaning according to which notions such as "proposition", "propositional content (of an attitude)" and so on are analysed in terms of the notion of a possible world. Thus a proposition is a set of possible worlds (the set of "those worlds in which the proposition is true"), a necessary truth is a sentence or proposition that is true in all possible worlds, and similarly for other such notions. This proposal for the analysis of propositional content and of intensional sentence operators and predicates is of a piece with the thesis that it is this very notion of propositional content — which identifies propositions with sets of possible worlds — that serves as basis for the analysis of "belief contexts" (and other "attitude contexts" for other types of attitudes): the complement clauses of verbs like *believe* and other attitudinal verbs contribute to the meaning of the whole their "propositional content" in the sense under discussion (i.e. the set of possible worlds which verify the complement clause).

It is an old and often repeated observation that this cannot be right [Turner, 1988]. The principle that propositions, in the present, intensional sense of the term according to which they are sets of worlds, are the "objects of belief" does not do justice to the form in which the content of what is believed is available to the believer. Suppose that Bill says that he believes that there are twice as many women in his class as men and that he has expressed his belief in these very words. There are innumerable ways of expressing this proposition — that the number of women in Bill's class is twice the number of men — in an intensionally equivalent way. Some such ways can be quite indirect, e.g. by restating the concept of one number being twice as large as another number in more esoteric mathematical terms, which require the know-how of an expert in number theory to be recognised as mathematically equivalent to the notion of multiplication by 2. And of course there is no limit to how abstruse the chosen formulations can be.[82] Most of these Bill — let us assume him to be a person of average mathematical knowledge and ability, though in the end the assumption matters little — will not recognise as expressing the belief to which he has just committed himself in the words mentioned. Yet they are all intensionally equivalent — i.e. they express the same propositional content, if propositional content is what the intensional approach makes it — as the words he has used himself. So if we take him by his word and attribute to him the belief which he has claimed for himself, then we are forced to say that Bill believes what each of the other sentences expresses too, notwithstanding his reluctance or refusal to accept them as true. In this manner the intensional approach calls into question one of the principal criteria that we use to determine what it

[82] A very modest step in this direction would be to say that any set containing the number of the men and closed under the operation of forming addition will contain the number of the women.

is that other people believe. (And mutatis mutandis for determining their other attitudes, such as desire, intention, etc.).

One reason why this is a fundamental problem is that it is directly connected with the question whether agents can arrive at new knowledge through ratiocination alone. We believe that it is one of the fundamental intuitions of the pure mathematician that it is possible to acquire new knowledge (and with it belief) through mathematical proof — that a mathematician who has established a surprising mathematical fact by finding a hard and non-obvious proof for it which reasons from mathematical axioms that every mathematician accepts and that he himself had been long acquainted with and never questioned, has discovered and established a new item of mathematical knowledge. However, if new knowledge can be gained through the transformation of information structures that were already there, without addition of any new information from outside, then evidently there is more to the form of information than the intensional approach allows for.

Accepting the verdict that seems to follow from these considerations is tantamount to condemning all intensional analyses of the propositional attitudes. This is a step that should not be taken lightly, for the intensional approach has proved immensely useful, and especially in the semantic analysis of natural language. It combines great simplicity with a degree of empirical adequacy which, although it could not be perfect (this is what the above reflections show beyond doubt), is nevertheless a striking advance over what came b efore it. That Montague Grammar has long been hailed — and still is by many, especially for its account of "intensional" contexts (attitude contexts prominently among them) — isn't due to collective confusion or bewitchment. Nevertheless, the step appears inevitable.

In fact, one could not hope for a *model-theoretic* approach towards the notions of meaning and inference to come much closer to a correct analysis of propositional attitudes and attitude ascription than the intensional approach actually does. For those distinctions which the theory of propositional attitudes needs but which the intensional approach cannot supply are not distinctions in truth conditions. They concern different ways in which the same truth conditions can be expressed. A theory which does justice to this aspect of the having and handling of information must therefore include a component which deals with the possibilities of transforming one representation of information into another one (which either expresses the very same information or some part of it), and as part of this addresses the question how hard or easy it is to carry out those transformations are. In other words, such a theory must include a *proof-theoretical* component.

On the face of it the hope that DRT could help us to develop such a theory is not unreasonable. For we have seen that the content representations which DRT makes available can distinguish in at least some cases between different expressions of the same propositional intension. This, one might say, is the true moral from a cognitive point of view of Partee's marble example (see (42), Section (3.2)), which played such a decisive part in demonstrating the importance of DRSs as a significant level of representation of linguistic information: "One of the ten marbles is not in the bag," and "Nine of the ten marbles are in the bag.", though expressing the same propositional intension, differ nevertheless in some way that has to do with their semantics, something which is captured by the DRSs for these two sentences and which turns out to be crucial if they are followed by a statement in which the pronoun "it" is intended to refer to the missing marble.

Dynamic Semantics succeeded in recasting the distinction between these two DRSs in the form of a kind of "refined intensionality", viz. by replacing the classical notion of a proposition — that of a set of possible worlds — by that of an information state — a set of pairs, with each pair consisting of a possible world and a "verifying embedding" which assigns objects to a certain set of discourse referents. (Definition (22), Section (3.2).) This refined intensionality concept, of which information states are the most salient representatives, is one step in the right direction — one step away from a classically intensional account of the attitudes and towards an account which pays due heed to issues of form and of transformation of forms through inference. But it is only one step, and a fairly minor one at that. The more serious obstacles to a model-theoretic account of the content of propositional attitudes and the semantics of attitude descriptions are cases with which this refined notion of intensionality cannot deal any better than the classical notion. As a rule these cases have nothing to do with the availability or non-availability of discourse referents which the refined notion is able to capture while the classical notion cannot.

Whether structural properties of DRSs other than what is contained in their main universes can be used to arrive at better approximations of intensionality is a question which cannot be answered here.[83] We doubt, however, that even if such other properties should prove to be cognitively significant, they could do more than give us what would still be only a partial solution of the intensionality problem. For there is one sense in which the intensional solution seems just right: once someone has been shown that two sentences are intensionally equivalent — i.e. that they are true in the same possible worlds — he simply can no longer sincerely profess belief in what the one sentence says and refuse to profess belief in what is said by the other; and likewise with other attitudinal modes, such as intending or desire.

In the model-theoretic semantics for L_{PA} we now proceed to present the problems which necessarily beset any version of the intensional approach towards the analysis of the attitudes have been set aside. Still, the semantics

[83] But compare for instance [Asher, 1986], where the formal similarity of DRSs is used to define a new notion of propositional identity which is much stricter than the notion of an information state (and thus a fortiori than the classical notion of a proposition as a set of possible worlds).

does take account of complexities illustrated by the Partee example, which means that we need at a minimum model-theoretic concepts such as that of an infromation state (Hence the spate of defintions of such notions in Section 3.2.) However, the notions we will actually need are more complicated yet. This is connected with the form of “naive realism about propositional attitudes that is adopted in our model theory. We assume that the information which renders attitudinal conditions true in a given model are psychological facts encoded in the model which pertain to the relevant attitude bearers at the relevant times. That is, we assume that each model $\mathcal{M}$ is equipped with a function $\mathrm{AS}_{\mathcal{M}}$ which assigns in each possible world w of $\mathcal{M}$ to each member a of a certain set CA_{W} of the universe of the model (intuitively: the Cognitive Agents of $\mathcal{M}$ in w) at each moment of time t belonging to a certain interval or set of intervals (the period(s) of consciousness of a in $\mathcal{M}$ in w) a certain object which identifies a’s mental state at the time in question. These objects are similar in structure to the Attitude Description Sets K which occupy the second argument position of the predicate Att. The values which $\mathrm{AS}_{\mathcal{M}}$ assigns to argument combinations w,a,t are not sets of pairs each consisting of a mode indicator and a DRS, but rather sets of pairs each consisting of a mode indicator and an “intensional object definable by a DRS”. In other words, the information about attitudinal states which is incorporated in the model $\mathcal{M}$ abstracts from the form of DRSs all but what is captured by an intensional semantics for DRSs.

What are the “intensional objects definable by a DRS” of which the last paragraph speaks? That will depend in the first place on what kind of intensional semantics we adopt. In the light of all that has been said about the importance of the role of discourse referents in an account of propositional attitudes the natural choice here is for the refined intensionality provided by information states — this, we take it, requires no further argumentation. However, it will not do to assume that the second members of pairs in $\mathrm{AS}_{\mathcal{M}}(\mathrm{w,a,t})$ are always simply information states. Recall that one of the points of ADSs K as attitudinal state descriptions was that they may contain *improper* DRSs — these, we saw, are the natural representations of attitudes which referentially depend on other attitudes that are part of the same mental state. An improper DRS, however, does not determine an information state on its own. By itself, it only defines an *information update* (or context update), and to get an information state out of this, the information update has to be combined with the information states determined by the attitudes on which it depends. Since the concept of referential dependence within a single attitudinal state is in principle recursive — one component K_1 of the state may referentially depend on another component K_2, which in its turn referentially depends on a component K_3, and so on — the intensional structure of an attitudinal complex can get quite involved. The definitions below, which build on those of an information state and a Context Change Potential as defined in Section (3.2), are designed to cope with this complication. In the interest of space we will give these definitions with a minimum of elucidation.[84]

The main problem that these definitions are designed to cope with are the referential dependencies of some components of a mental state on others. Suppose that the component K_1 of a given mental state depends on the components $K_{2,1}$ and $K_{2,2}$ and on no others, and that of these $K_{2,1}$ depends in turn on the components $K_{3,1}$ and $K_{3,2}$ and no others; and furthermore that $K_{2,2}$, $K_{3,1}$ and $K_{3,2}$ do not depend on any other components. Let us also assume that none of these components are internal anchors. $K_{2,2}$, $K_{3,1}$ and $K_{3,2}$ will be proper DRSs and thus each determine an information state (with respect to a given model $\mathcal{M}$, world w of $\mathcal{M}$ and time t_0 of $\mathcal{M}$). (They also define, as explained in Section (3.2), a regular total CCP; these CCPs stand to the information states in the relation stated there.) The other two components, K_1 and $K_{2,1}$, will be improper DRSs and thus determine non-total CCPs relative to $\mathcal{M}$, and no information states. Note however that the CCP $\mathcal{J}_{2,1}$ determined by $K_{2,1}$, will be defined for the merge $\mathcal{I}_{3,1} \uplus \mathcal{I}_{3,2}$ of the information states $\mathcal{I}_{3,1}$, $\mathcal{I}_{3,2}$ determined by $K_{3,1}$ and $K_{3,2}$. (The reason is that according to our assumptions the bases of $\mathcal{I}_{3,1}$ and $\mathcal{I}_{3,2}$ together cover the set of free discourse referents of $K_{2,1}$.) Let $\mathcal{I}_{2,1}$ be the result of applying $\mathcal{J}_{2,1}$ to $\mathcal{I}_{3,1} \uplus \mathcal{I}_{3,2}$, i.e. $\mathcal{I}_{2,1} = \mathcal{J}_{2,1}(\mathcal{I}_{3,1} \uplus \mathcal{I}_{3,2})$. The same considerations lead us to conclude that the CCP defined by K_1 will be defined for the merge $\mathcal{I}_{2,1} \uplus \mathcal{I}_{2,2}$ (where $\mathcal{I}_{2,2}$ is the information state determined by $K_{2,2}$); so we can also associate an information state, viz. $\mathcal{J}_1(\mathcal{I}_{2,1} \uplus \mathcal{I}_{2,2})$, with K_1.

The moral of this should be clear: By themselves the dependent components of a complex attitudinal state do not define information states; but they do so when combined with the information states determined by the components upon which they depend — provided those determine information states, something they will do so long as this is true for the components that they depend on — and so on, all the way down. This of course presupposes that by going all the way down one comes, no matter how one goes, to a well-defined end. In other words, the referential dependence relation between components of a mental state should be well-founded. So we will assume that the attitude description sets K to which the semantics described in this section assigns intuitively acceptable model-theoretic interpretations are all well-founded in the sense that the transitive closure $\prec_K$ of the following relation $\prec$ between the DRS components K_1 and K_2 of an ADS K is well-founded: $K_1 \prec K_2$ iff there is a discourse referent x which occurs free in K_2 and belongs to the universe of K_1. We will from now on assume that we are dealing only with ADSs which satisfy this well-foundedness constraint.

In addition we restrict attention to ADSs K which are "proper over all" in that for each $\langle \text{MOD, K} \rangle \in K$ the set Fr(K) of free discourse referents

[84] For more extensive comments see [Kamp, 2003].

of K is included in the union of the universes of DRSs occurring in pairs $\langle \mathrm{MOD}', \mathrm{K}' \rangle \in K$ such that $\mathrm{K}' \prec_K \mathrm{K}$:

$$\mathrm{Fr(K)} \subseteq \cup\{\mathrm{U}_{\mathrm{K}'} \mid (\exists\ \mathrm{MOD}')\langle \mathrm{MOD}', \mathrm{K}' \rangle \in K \wedge \mathrm{K}' \prec_K \mathrm{K}\}$$

Given these assumptions about ADSs, it will be possible, given a model $\mathcal{M}$, a world w in $\mathcal{M}$ and an instant of time t of $\mathcal{M}$, to associate with each component DRS K of an ADS K an information state $[\![\mathrm{K}]\!]^s_{\mathcal{M},\mathrm{w},\mathrm{t},K}$. We cannot show this yet, since we haven't made any commitments on the form of the information given by the function $\mathrm{AS}_{\mathcal{M}}$. But we can illustrate the general idea for simple ADSs K, in which all content specifications K (i.e. all second components of members $\langle \mathrm{MOD,K} \rangle$ of K) are DRSs of the underlying DRS language L, in which the predicate Att does not occur. For the evaluation of such K the function $\mathrm{AS}_{\mathcal{M}}$ plays no role, so we may assume that $\mathcal{M}$ is a model for the underlying language L. First suppose that K is a content specification which has no predecessors in the order $\prec_K$. Then $[\![\mathrm{K}]\!]^s_{\mathcal{M},\mathrm{w},\mathrm{t},K}$ is simply the information state determined by K in $\mathcal{M}$ in w at t. Secondly, suppose that we have determined information states $[\![\mathrm{K}']\!]^s_{\mathcal{M},\mathrm{w},\mathrm{t},K}$ for all content specifications K′ in K such that $\mathrm{K}' \prec_K \mathrm{K}$. By assumption the CCP $\mathcal{J}(\mathrm{K},\mathcal{M},\mathrm{w},\mathrm{t},K)$ determined by K in $\mathcal{M}$ in w at t will be defined for the merge of these information states: $\underline{\cup}\ \{\ [\![\mathrm{K}']\!]^s_{\mathcal{M},\mathrm{w},\mathrm{t},K} \colon \mathrm{K}' \prec_K \mathrm{K}\ \}$.

We noted earlier that a model-theoretic analysis of when the attitude descriptions provided by ADSs are correct implies that any model $\mathcal{M}$ must contain information about the actual attitudinal state of an agent a in a world w at a time t in terms of which ADSs can be evaluated; and we assumed that this information is supplied by the function $\mathrm{AS}_{\mathcal{M}}$. We must now decide in which form this information is given. We will opt for a form which makes the evaluation of ADSs in models a comparatively straightforward matter.

We proceed in two steps. We first define the notion of a P(otential) I(nformation) S(tate) B(ased) A(ttitudinal) S(tate description), and then narrow this concept down further to that of an I(nformation) S(tate) B(ased) A(ttitudinal) S(tate description).

DEFINITION 63. Let $\mathcal{M}$ be a model and let $\mathcal{J}$, $\mathcal{J}_1$, $\mathcal{J}_2$, $\mathcal{J}'$ be CCPs:

(i) A PISBAS *relative to* $\mathcal{M}$ is any set of pairs $\langle \mathrm{MOD}, \mathcal{J}\ \rangle$, with MOD a mode indicator and $\mathcal{J}$ a regular CCP relative to $\mathcal{M}$.

(ii) Let J be a PISBAS relative to $\mathcal{M}$. Let $\prec_J$ be the transitive closure of the relation $\prec$ between the members of J defined by: $\mathcal{J}_1 \prec \mathcal{J}_2$ iff there is a discourse referent x which belongs to $\mathrm{PRES}(\mathcal{J}_2)$ and to a base of $\mathcal{J}_1$.

(iii) We say that a PISBAS relative to $\mathcal{M}$ is an ISBAS *relative to* $\mathcal{M}$ iff (i) $\prec_J$ is well-founded and (ii) it is possible to assign, by induction along $\prec_J$, to each CCP $\mathcal{J}$ occurring in J an information state $\mathrm{I}(\mathcal{J})$ as follows: (a) Suppose that $\mathcal{J}$ has no predecessors according to $\prec_J$. Then $\mathcal{J}$ is a total CCP and the associated information state $\mathrm{I}(\mathcal{J})$ is defined as $\mathcal{J}(\Lambda)$. (b) Suppose that for all $\mathcal{J}'$ occurring in J such that $\mathcal{J}' \prec_J \mathcal{J}$, $\mathrm{I}(\mathcal{J}')$ has been defined. Then $\mathcal{J}$ is defined on $\uplus\{\mathrm{I}(\mathcal{J}') \mid \mathcal{J}' \prec_J \mathcal{J}\}$ and $\mathrm{I}(\mathcal{J}) = \mathcal{J}(\uplus\{\mathrm{I}(\mathcal{J}') \mid \mathcal{J}' \prec_J \mathcal{J}\})$.

The idea behind the definition of a PISBAS is that of a structure that is essentially like that of an ASD, except that the DRSs which form the second components of the pairs $\langle$MOD, K$\rangle$ which occur as elements of ASDs are replaced by intensional objects (relative to the given model $\mathcal{M}$) of the sorts that DRSs can be used to describe. Since the DRSs occurring in ASDs are sometimes improper, these intensional objects cannot always be information states; in general they will have to be CCPs. However, when a PISBAS is an ISBAS, each of these CCPs is, roughly speaking, defined on the merge of information states that can be associated with all the CCPs on which it "referentially depends": the $\mathrm{I}(\mathcal{J}')$ such that $\mathcal{J}' \prec_J \mathcal{J}$ jointly fulfill the presupposition of the CCP $\mathcal{J}$, so that application of $\mathcal{J}$ to their merge gives a well-defined information state, viz. $\mathrm{I}(\mathcal{J})$. The concept of an ISBAS thus captures the idea that the contents of propositional attitudes which make up a complex attitudinal state may depend on other attitudes in the state, but that in possible worlds where the propositional contents of these other attitudes are true, the dependent attitude has a well-defined propositional content. In restricting attention to ISBASs as the possible values of the function $\mathrm{AS}_{\mathcal{M}}$, and thus as the only possible characterisations of complex attitudinal states (of a person a in a world w at a time t) we thus impose a certain coherence condition on the mental states that, according to our model theory,it is possible for a cognitive agent to be in.

There is at least one further constraint that it seems reasonable to impose on the possible values of $\mathrm{AS}_{\mathcal{M}}$. In our first version of the stamp example (see (211) ff.) the desire and intention referentially depend on the belief. Such a state of affairs seems quite possible intuitively: You have a belief to the effect that a certain thing exists and then form regarding the thing you believe to exist a certain desire and/or intention. But can a belief referentially depend on a desire or an intention? We think not. On the face of it this might perhaps seem like a possibility — something of the order of wishful thinking, not to be recommended perhaps, but a cognitive possibility even so. When we look more closely, however, we realise that wishful thinking is really something else. In wishful thinking a desire may be the irrational and unjustifiable cause of a belief. But the belief won't be referentially grounded in the desire in the way in which we have seen that a desire can be referentially dependent on a belief.[85]

To capture this additional constraint we need to specify the *Attitudinal Hierarchy.* This is a partial order $<_{\mathrm{MOD}}$ between mode indicators; $\mathrm{MOD}_1 <_{\mathrm{MOD}} \mathrm{MOD}_2$ means that an attitude of mode MOD_2 may referentially depend on one of mode MOD_1. With only the three mode indicators BEL, DES, INT we would, in the light of the remarks above, assume that BEL $<_{\mathrm{MOD}}$ DES and BEL $<_{\mathrm{MOD}}$ INT, as well as BEL $<_{\mathrm{MOD}}$ BEL (whereas the relations DES $<_{\mathrm{MOD}}$ BEL and INT $<_{\mathrm{MOD}}$ BEL never hold). Whether $<_{\mathrm{MOD}}$ should be assumed to hold also between DES and INT, however, or between INT and DES, is a more delicate question. We will not try to solve these here. When further mode indicators are added, the Attitudinal Hierarchy must be extended. For instance, addition to the set {BEL,DES,INT} of the mode indicator ANCH, as shown in (210), comes with an extension of the relation $<_{\mathrm{MOD}}$ with all pairs ⟨ANCH, MOD⟩ where MOD is any one of the indicators ANCH, BEL, DES, INT. More generally, richer and more refined mode indicator classifications each come with their own Attitudinal Hierarchy, and each raises its own problems about what that hierarchy is like.

We have now laid the groundwork for the truth definition that is the central purpose of this subsection. We make the general assumption that for any model $\mathcal{M}$, $\mathrm{AS}_{\mathcal{M}}$(a,w,t), if defined, is an ISBAS relative to $\mathcal{M}$.[86] To define truth (=proper embeddability) of DRSs of our extended language L_{PA} into such models we proceed in three steps. In the remainder of this subsection we give the truth definition for that sublanguage of L_{PA} in which (i) ADSs contain no internal anchors (and in which, consequently, there are no external anchors either; thus the third argument of Att will always be the empty set and we can treat Att as a 2-place predicate); and (ii) in which there are no occurrences of i and of n within the scope of Att. In the next subsection, 5.3, we deal with the reference conditions of i and n, and in subsection 5.3 with the full language L_{PA}.

The remaining task of this subsection is to state the verification conditions for DRS-conditions of the form "s:Att(a,K)". In outline it should be clear what these conditions ought to be: an embedding function f verifies the condition in $\mathcal{M}$ in w iff for each time t within the duration of f(s), K is a correct description of $\mathrm{AS}_{\mathcal{M}}$($f$(a),w,t). Given our decision about the values of AS, it should also be roughly clear how we should interpret the phrase "K is a correct description of $\mathrm{AS}_{\mathcal{M}}$($f$(a),w,t)": There must exist a map H

[85]The belief could become in its turn the basis for the emergence of a further desire with a content which referentially depends on the belief. But the referential dependence will still be this way round, not of the belief on the first desire.

[86]Note well that in doing so we adopt the intensional perspective which we criticised because of its inability to deal with the logical equivalence problem. But as we already noted a refined intensional treatment of propositional attitudes is the best we can do within a framework that is purely model-theoretic.

from K to $\mathrm{AS}_{\mathcal{M}}(f(\mathrm{a}),\mathrm{w},\mathrm{t})$ such that for each $\langle$MOD,K$\rangle \in K$, the Mode of H($\langle$MOD,K $\rangle$) matches that of $\langle$MOD,K$\rangle$ and the content of H($\langle$MOD,K$\rangle$) matches the content of $\langle$MOD,K $\rangle$. But what is matching here? It is not, we contend, quite the same in the two cases. Matching of Mode should (at least for the extremely simple Mode Indicator system used here) be just what the term suggests, viz. identity: if H($\langle$MOD,K $\rangle$) = $\langle$MOD$'$,J$\rangle$, then it must be the case that MOD$'$ = MOD. In connection with content, however, identity does not seem the right way to define matching. We normally regard an attitude description as correct even if it is not complete. This fact is particularly striking for the attitudinal modes of desire and intention. We can truthfully describe Mary as wanting to marry a Swede not only when her goal is as unspecific as simply "marrying a Swede" (which it it would be unlikely to be), but also (more plausibly) when her idea of a suitable husband goes well beyond that: what she wants is not just any Swede, but one who is tall, blond, blue-eyed, and (of course) handsome, dashing and considerate. In other words, the content of her actual desire may be much richer than the description which we give of it. On the other hand, in order that the description is to count as correct, it must subsume the actual content.

For belief the argument that content matching should be defined as logical entailment of the described belief by the one actually held according to $\mathrm{AS}_{\mathcal{M}}$ is not quite the same as for desire or intention, and arguably it is somewhat less persuasive. According to our own intuitions, however, belief attribution also obeys the principle of content subsumption, so we will handle matching for description components of the form $\langle$BEL,K$\rangle$ in the same way as those of the forms $\langle$DES,K$\rangle$ and $\langle$ INT,K$\rangle$.[87]

How do we capture subsumption of K by H($\langle$MOD,K$\rangle$)? Suppose that H($\langle$MOD,K$\rangle$) = $\langle$MOD,$\mathcal{J}$ $\rangle$. In view of the assumptions which we have been making about ADSs on the one hand and about their model-theoretic counterparts, ISBASs, on the other, it might be thought that subsumption can be stated straightforwardly: the information state $[\![\mathrm{K}]\!]^{s}_{\mathcal{M},\mathrm{w},\mathrm{t},K}$ must be entailed, in the sense of entailment that is appropriate for information states, by the information state I($\mathcal{J}$) which, we have seen, can be associated within the ISBAS $\mathrm{AS}_{\mathcal{M}}(f(\mathrm{a}),\mathrm{w},\mathrm{t})$ with the CCP $\mathcal{J}$; that is, $[\![\mathrm{K}]\!]^{s}_{\mathcal{M},\mathrm{w},\mathrm{t},K} \preceq \mathrm{I}(\mathcal{M})$.

But there is one further snag here: the discourse rerferents occurring in the ADS K need not be the same as those occurring in the bases of the CCPs of the ISBAS. Some "renaming of variables" is needed in order to

[87]Note well that the subsumption principle does not hold for all attitudinal modes. It doesn't hold, for instance, for doubt, or for "wondering", the attitude which an agent a entertains vis-a-vis a proposition p when A is unsure whether p is true and wonders whether or not it is. For attitude descriptions with richer mode repertoires the verification condition below will therefore have to be more complicated than it is for the restricted set {BEL,DES,INT}.

make sure that the information states $[\![K]\!]^s_{\mathcal{M},w,t,K}$ and $I(\mathcal{J})$ can be related in the right way. This clearly requires that the bases of the first be included in those of the second. But of course this need not be the case, even if the attitude description provided by K is intuitively correct. For the discourse referents chosen in the actual description which K provides can be chosen freely, and will in general stand in no relation to those of the ISBAS. There are two ways to get rid of this discrepancy — either we rename the ADS or we rename the ISBAS.

For reasons which will become transparent later on we prefer the first of these options. This however runs into another difficulty, which is also connected with the formal identity of discourse referents. The ADS that needs evaluation may be part of a larger DRS in which discourse referents occur "higher up" which happen to be part also of the ISBAS. Renaming bound discourse referents from the ADS into such discourse referents could wreak havoc with the proper functioning of the truth definition and should be avoided. We eliminate this danger once and for all by assuming that the discourse referents occurring in ISBASs (including in particular all those which occur as values of the function AS_M are entirely disjoint from those which belong to the language L_PA.

Suppose that r is a 1-1 map from the set of discourse referents occurring in the ADS K onto some other set of discourse referents. Then the *alphabetic variant of* K *determined by* r is the set of all pairs $\langle$MOD,r(K)$\rangle$ such that $\langle$MOD,K$\rangle$ belongs to K together with the pairs $\langle$[ANCH,r(x)],r(K)$\rangle$ such that $\langle$[ANCH,x].K$\rangle$ belongs to K. r(K) is the DRS obtained by replacing each discourse referent x occurring in K throughout K by r(x).

At last we have all the pieces we need to state the verification conditions for DRS conditions of the form s:Att(a,K). We get:

DEFINITION 64.

$f \models_{\mathcal{M},w}$ s:Att(a,K) iff there exists (i) a renaming function r such that Dom(r) consists of the discourse referents occurring in K and (ii) a function H with Dom(H) = r(K) such that (a) H($\langle$MOD,K$\rangle$) is of the form $\langle$MOD,$\mathcal{J}$ $\rangle$, (b) for all t $\in$ dur(f(s)) and each $\langle$MOD,K$\rangle \in$ r(K) H($\langle$MOD,K$\rangle$) belongs to $AS_{\mathcal{M}}(f(a),w,t)$ and (c) for each $\langle$MOD,K$\rangle \in$ r(K), $[\![K]\!]^s_{\mathcal{M},w,t,K} \preceq I(\mathcal{J})$, where $I(\mathcal{J})$ is the information state determined within $AS_{\mathcal{M}}(f(a),w,t)$ by the CCP $\mathcal{J}$ of H($\langle$MOD,K$\rangle$).

The Indexical Discourse Referents i and n

We need two further specifications, concerning the indexical discourse referents i and n. i, we stipulated, only occurs within the scope of Att. And there it always represents the self of the cognitive agent which appears as first argument of Att. The matter of its interpretation is slightly more complicated, however, than this informal description may suggest, for some occurrences of Att may occur within the scope of others. For instance, we can express in our DRS language the statement that Bill thinks that Mary thought that she was clever. (More precisely, that Mary had a thought which she herself might have expressed as "I am clever."). The condition expressing this is given in (217)

(217)
$$\begin{array}{|c|}\hline \mathrm{t}_1\ \ \mathrm{s}_1\ \ \mathrm{b} \\ \mathrm{t}_1 = \mathrm{n}\ \ \mathrm{n} \subseteq \mathrm{s}_1\ \ \mathrm{Bill(b)} \\ \mathrm{s}_1{:}\mathrm{Att(b,} \left\{ \left\langle \mathrm{BEL,} \begin{array}{|c|}\hline \mathrm{t}_2\ \ \mathrm{s}_2\ \ \mathrm{m} \\ \mathrm{t}_2 < \mathrm{n}\ \ \mathrm{t}_2 \subseteq s_2\ \ \mathrm{Mary(m)} \\ \mathrm{s}_2{:}\mathrm{Att(m,} \left\{ \left\langle \mathrm{BEL,} \begin{array}{|c|}\hline \mathrm{s}_3 \\ \mathrm{n} \subseteq \mathrm{s}_3 \\ \mathrm{s}_3{:}\ \mathrm{clever(i)} \\ \hline \end{array} \right\rangle \right\}) \\ \hline \end{array} \right\rangle \right\}) \\ \hline \end{array}$$

(217) has two occurrences of Att, one within the scope of the other, the first argument of the outer occurrence is the discourse referent b representing Bill, that of the inner occurrence the discourse referent m representing Mary. Clearly it is Mary whose self the occurrence of i in (217) is meant to represent. The general principle should be clear from this example: an occurrence of i represents the self of the first argument of the nearest occurrence of Att one encounters when going upwards from that occurrence in the structure of the DRS.

Formally this means, first, that the entity denoted by an occurrence of i must be evaluated within the context of this occurrence - i.e. with respect to the DRS K which contains it. And because K may well contain several occurrences of i, we need some device to distinguish these. To this end we assume that the different occurrences of i in K are indexed and use the symbol "$\mathrm{i}_{(\mathrm{j,K})}$" to refer to the j-th of these occurrences.

Secondly, the denotation of $\mathrm{i}_{(\mathrm{j,K})}$ is determined by Definition (65)

DEFINITION 65. $[\![\mathrm{i}_{(\mathrm{j,K})}]\!]_{\mathcal{M},\mathrm{w},f} = f(\mathrm{a})$ where a is the discourse referent occupying the first argument slot of that occurrence of Att in K which contains $\mathrm{i}_{(\mathrm{j,K})}$ in its scope and is within the scope of all other occurrences of Att in K with this property.[88]

[88]Denotation clauses for singular terms like that in (65) haven't been considered so far, and may seem at variance with the way in which verification and truth definitions are usually formulated for DRS languages. However, the change is only a slight one. Even at this point there are only two kinds of terms to be considered,(i) "ordinary" discourse referents, any occurrences of which in proper DRSs are bound by an occurrence of the discourse referent in some DRS universe, and (ii) the two indexical discourse referents n and i. The former discourse referents will, in any normal evaluation of a proper DRS, already be in the domain of the embedding function under consideration when the question arises whether the function verifies a condition which contains such

N.B. The way in which DRSs and their parts are semantically evaluated guarantees that by the time we "get to the given occurrence $\mathrm{i}_{(\mathrm{j,K})}$ of i in K", the embedding function f will be defined for the relevant argument a.

The interpretation of n is determined by much the same principles as that of i: when n occurs within the scope of an occurrence of Att, then it is intended to represent the "present" of the represented thought, i.e. as representing the present from the perspective of the thinker at the time when he had that thought. That is, the occurrence should be interpreted as referring to the very same time as that of the state s characterised as "s:Att(a,K)", where the given occurrence of n is somewhere in K. Consider for instance the occurrence of n in (217) as part of the condition "$\mathrm{n} \subseteq \mathrm{s}_3$". This occurrence marks the time at which Mary has the thought which according to (217) Bill attributes to her. This is the time represented by t_2, which according to what (217) says is in the past of the time represented by the occurrence of n in the condition "$\mathrm{t}_2 < \mathrm{n}$". And that time, the time of the thought of Bill, and thus of the corresponding state s_1, is one which includes the utterance time of the entire statement represented by (217). It is easy to see that each of these occurrences of n is made to refer to the intuitively right time if we stipulate that the value assigned to an occurrence of n in K by an embedding function f is equal to dur(f(s)), where s is the state discourse referent such that the occurrence of n is in the condition s: Att(α,K') and where moreover this is the nearest condition of this form containing that occurrence.

One difference between i and n is that n is also allowed to occur outside the scope of Att. In those cases it refers to the utterance time of the represented statement. Thus the interpretation clause for n divides into two parts. (In analogy with our convention for i, we denote particular occurrences of n in K as "$\mathrm{n}_{(\mathrm{j,K})}$".)

DEFINITION 66.

(i) Suppose that the occurrence $\mathrm{n}_{(\mathrm{j,K})}$ of n in K is within some condition of the form s:Att(α,K'). Then

$[\![\mathrm{n}_{(\mathrm{j,K})}]\!]_{\mathcal{M},\mathrm{w},f} = \mathrm{dur}(f(\mathrm{s}_0))$, where s_0 is the discourse referent such that $\mathrm{n}_{(\mathrm{j,K})}$ occurs in s_0:Att(α,K') in K and s_0:Att(α,K') is within the

discourse referents as arguments, and the values of these discourse referents will then be whatever this function assigns to them; and the values which ambedding functions assign to i and n are determined once and for all by 66 and 68 below. This fixes the values of $[\![\alpha]\!]_{\mathcal{M},\mathrm{w}}$ for all relevant cases — both when α is an ordinary discourse referent and when it is i or n. The verification clauses for atomic conditions will now refer to the values of their argument terms. For instance, the clause for an atomic condition $\mathrm{P}(\alpha_1,\ldots,\alpha_n)$ can now be stated by referring to the values (under the embedding function f in question) of their argument terms:

$$f \models_{\mathcal{M},\mathrm{w}} \mathrm{P}(\alpha_1,\ldots,\alpha_n) \text{ iff } \langle\, [\![\alpha_1]\!]_{\mathcal{M},\mathrm{w},f},\ldots,[\![\alpha_n]\!]_{\mathcal{M},\mathrm{w},f} \,\rangle \in \mathrm{F}_{\mathcal{M},\mathrm{w}}(\mathrm{P}).$$

scope of all other conditions of this form which contain $n_{(j,K)}$.

(ii) Suppose that the occurrence $n_{(j,K)}$ of n in K is not within any condition of the form s: Att(α,K'). Then

$[\![n_{(j,K)}]\!]_{\mathcal{M},w,f}$ = the "utterance time of the represented utterance.

Semantics for Anchored Representations

The verification definition (64) only covers representations in which all discourse referents are unanchored. When anchored discourse referents are taken into account, matters get a little more complicated. First, we now must distinguish between wide content and narrow content. In the case of wide content, the internal and external anchors play a part in the verification conditions, in the case of narrow content they do not.

We start with wide content. There are two complications which do not arise with anchor-free representations. First, when a discourse referent x which is internally anchored in an ADS K has an external anchor x′, then each DRS K such that ⟨MOD,K⟩ belongs to K and in which x occurs should be seen as expressing a proposition that is singular with respect to the value of x′. More precisely, in the context of evaluating the condition "s:Att(a,K,EA)" in $\mathcal{M}$ in w at t under f the proposition expressed by K in $\mathcal{M}$ relative to f should be singular with respect to f(x′). (Note that if ⟨x,x′ ⟩ ∈ EA, then x′ occurs free in "s:Att(a,K,EA)"; so if evaluation of "s:Att(a,K,EA)" in $\mathcal{M}$ in w at t under f arises in the context of evaluating a proper DRS of L_{PA} in which the condition occurs, then x′ will be in the Domain of f.) We achieve singularity of the proposition expressed by K with respect to all the internally and externally anchored discourse referents occurring in K if we evaluate the proposition expressed by K not with respect to f but with respect to the extension $f \cup (\mathrm{EA} \circ f)$ of f which has each of these discourse referents x in its domain and assigns to x the value that f assigns to x′.

The second desideratum for the verification condition for "s:Att(a,K,EA)" is that verification is undefined when K contains discourse referents which are internally but not externally anchored. There are various ways to achieve this. A very simple one is to remove the internal anchors of such discourse referents from K. This will in particular have the effect that occurrences of the discourse referents whose anchors have been removed in other components of K will not be declared (i.e. they won't belong to any DRS universe). As always this causes indeterminacy of verification for any atomic condition which contains such a discourse referent as argument. This will then also entail indeterminacy of the verification condition for "s:Att(a,K,EA)".

To implement this idea we must form, given an ADS K and an external anchor EA, the *Reduction of* K *with respect to* EA, Red(K,EA). This is the structure which we get by removing all internal anchors in K which aren't justified by EA, i.e. all internal anchors for discourse referents which do not occur in the Domain of EA:

DEFINITION 67. $\mathrm{Red}(K,\mathrm{EA}) := K \setminus \{\langle[\mathrm{ANCH,x}],\mathrm{K}\rangle \mid \langle[\mathrm{ANCH,x}],\mathrm{K}\rangle \in K \wedge \neg\, (\exists\, \mathrm{x}')\langle \mathrm{x},\mathrm{x}'\, \rangle \in \mathrm{EA}\, \}$

N.B. Evidently. if all internally anchored discourse referents of K are externally anchored by EA, then $\mathrm{Red}(K,\mathrm{EA}) = K$.

We are now ready to state the generalisation of Definition (64) in the sense of wide content:

DEFINITION 68. $f \models_{\mathcal{M},\mathrm{w}}$ s:Att(a,K,EA) iff for all t $\in$ dur(f(s)) there exists a function H from Red(K,EA) into $\mathrm{AS}_{\mathcal{M}}(f(\mathrm{a}),\mathrm{w},\mathrm{t})$ such that for each $\langle\mathrm{MOD,K}\rangle \in \mathrm{Red}(K,\mathrm{EA})$, $[\![\mathrm{K}]\!]^s_{\mathcal{M},\mathrm{w},f\,\cup\,(\mathrm{EA}\,\circ\, f),K} \preceq \mathrm{I}(\mathcal{J})$, where $\mathrm{I}(\mathcal{J})$ is the information state determined within $\mathrm{AS}_{\mathcal{M}}(f(\mathrm{a}),\mathrm{w},\mathrm{t})$ by the CCP $\mathcal{J}$ of $\mathrm{H}(\langle\mathrm{MOD,K}\rangle)$.

Our last task in this section is to define the verification conditions of "s:Att(a,K, EA)" in the sense of narrow content. Informally speaking this amounts to ignoring the external anchor EA and treating internally anchored discourse referents of K "existentially". Existential interpretation of the internally anchored discourse referents can be accomplished in more than one way, with slightly different effects. One of them is to treat the internal anchors as "de dicto beliefs", i.e. to replace each internal anchor $\langle$ [ANCH,x],K$\rangle$ in K by $\langle$BEL,K$\rangle$, and that is the one we adopt. To this end we define, for arbitrary ADS K:

DEFINITION 69. $\mathrm{NC}(K) = (K \setminus \{\langle[\mathrm{ANCH,x}],\mathrm{K}\rangle \mid \langle[\mathrm{ANCH,x}],\mathrm{K}\,\rangle \in K\,\}) \cup \{\langle[\mathrm{BEL,K}\rangle\!: \langle[\mathrm{ANCH,x}],\mathrm{K}\rangle \in K\,\}$

The narrow content verification of "s:Att(a,K, EA)" can now be defined as the verification of the condition "s:Att(a,NC(K))" in the sense of Definition (64).

We have argued that being in a mental state involving internally anchored discourse referents which lack an external anchor is being in a state involving unjustified presuppositions. So at least those attitudes that are part of the state and which are directly affected by the presupposition failure fail to determine well-defined propositions. Yet, we noted, the unjustified internal anchors are connected with existential beliefs whose truth conditions are well-defined, but false. In the light of the developments in this section it seems plausible that these remarks can now be made more explicit via the notion of narrow content: Given an ADS K, we obtain the associated beliefs by passing from K to NC(K).

Whether this gives us precisely what we want isn't altogether clear. For it isn't clear that the associated beliefs will necessarily be false. Such a belief, associated with an unjustified internal anchor for the discourse referent x, might come out true if there were an object satisfying the anchor's DRS (which is also the DRS of the belief which replaces the anchor in NC(K)), even though there was nothing to cause the introduction of x. Whether this is a genuine possibility depends on detailed assumptions about the conditions imposed by internal anchors. This is a matter that requires careful discussion and one that we decided to set aside in this survey.

There is however another way of associating beliefs with unjustified internal anchors. It involves a form of reflection — a thought process in which the agent reflects on his own thoughts, thereby making these into the subjects of further thoughts. The simplest form of reflection consists of nothing more than being aware that one has the thoughts one has. Reflection of this kind is possible, we take it, not only in relation to single attitudes but also to attitude complexes. Within the formalism developed in thi section the capacity of self-reflection comes to this: We assume that whenever an agent A is in a mental state that can be described by means of an ADS K, A is in a position to form beliefs of the form (218)

$$(218)\ \left\langle \text{BEL}, \boxed{\begin{array}{c} \text{s} \ \ \text{x}'_1 \ldots \text{x}'_n \\ \hline \text{n} \subseteq \text{s} \\ \text{s:Att(i,}K\text{,EA)} \end{array}} \right\rangle$$

(Here EA is $\{\langle \text{x}_1,\text{x}'_1 \rangle, ..., \langle \text{x}_n,\text{x}'_n \rangle\}$, with$\text{x}_1, \ldots, \text{x}_n$ the discourse referents with internal anchors in K.) [89]

While most beliefs which result from this type of reflection are true in virtue of the very fact that the agent does have the attitudes which (218) says he believes he has, this is not so for those instances of (218) in which K involves unjustified internal anchors, anchors for which there is no corresponding external anchor. For the lack of an external anchor is precisely what (218) denies: cases of unjustified internal anchors are cases where the agent is mistaken about what attitudes he has.

It is also reasonable to assume that reflection can target internal anchors by themselves. (219) gives the belief resulting from such a reflection:

$$(219)\ \left\langle \text{BEL}, \boxed{\begin{array}{c} \text{s} \ \ \text{x}' \\ \hline \text{n} \subseteq \text{s} \\ \text{s:Att(i,}\{\langle\ [\text{ANCH,x}]\text{,K}\ \rangle\}\text{, }\{\langle \text{x,x}'\ \rangle\}\ ) \end{array}} \right\rangle$$

where $\langle$[ANCH,x],K$\rangle$ is a correct description of one of the internal anchors belonging to A's attitudinal state at the time in question. For any internal anchor $\langle$[ANCH,x],K$\rangle$ of an attitudinal state described by K the belief represented in (219) is necessarily false if $\langle$[ANCH,x],K$\rangle$ is unjustified. Indeed,

[89] In addition, one might consider products of self-reflection beliefs which attribute to the x'_i some or all of the properties that are specified in the internal anchor of x_i. But these aren't needed for the present consideration.

the representations in (219) seem to capture exactly the idea of the false existential beliefs lurking behind defective attitudes de re.

5.4 Construction of Representations of Attitude Attributing Sentences and Texts

The formalism described in Section (5.2) has considerable flexibility. On the one hand it allows us to represent not only referentially connected attitudinal complexes, but also successions of these in time; thus it affords representation of attitudinal change, and not just static representations of attitudinal states at one given time. On the other hand, the formalism allows for the representation of thoughts whose content is itself an attitude attribution (either to someone else or to oneself, as when one reflects on one's own thoughts). Since this representational device is recursive, it allows also for thoughts that are attributions of attributions — as for instance when I wonder what you may be thinking about me — and so on. All these different aspects are important in a wide range of applications, and in particular in the description of the attitudinal states of participants in a conversation which arise through the verbal exchanges between them and guide the successive utterances through which the conversation progresses. (Of special importance in connection with the representation of conversation are mutliply iterated attitude attributions (of the type "You think that I think that you think that ...". such attributions play an important part in human interaction generally, and they are an almost inariable by-product of what happens when people talk face-to-face: "I have just said this and you know that I have and you know that I am aware that you know that I ...".)

To construct representations in which the various devices of our formalism are instantiated we need an extension of the DRS construction algorithm. In fact, the problem that many of the intended applications present is that we do not only need additional DRS construction rules to supplement the construction algorithms for the underlying DRS language (i.e.rules which extend the construction algorithm for the underlying language L to one for the full language L_{PA}); we also need rules that apply to settings not considered hitherto, such as that of a conversation in which speakers take turns. This second problem is a major one in its own right, which should be addressed in some other context. In this section we will only be concerned with the first, that of extending the text processing construction algorithms discussed in earlier sections to algorithms that can handle the problems of those sentences by means of which attitude attributions are made.

Of such attitude-attributing sentences we will only consider a very small sample, in which the vehicle for attitude attribution is an "attitude attributing verb". Examples of such verbs are *believe, know, hope, want, desire, re-*

gret, ... [90] Moreover, we will only look at a very small number of examples here, emphasising the problems which an extended construction algorithm will have to tackle. Our aim will be to bring to light the special problems which will have to be tackled when the construction algorithm is extended so that it covers attitude-attributing sentences of unrestricted form. But we will only give informal hints of how the solutions might go, sketching sonme of the additional construction principles that will be needed but without stating an extension explicitly. For further details we must refer the reader to [Hans and Reyle,].

A first and Simple Case of Interpretation Using Secondary Context

We start with an example that may be familiar to many. It was first discussed by Stalnaker in [Stalnaker, 1988] (See also the comments [Kamp, 1988] following this paper) It consists of two sentences:

(220) Phoebe believes that a man has broken into her garden. She thinks that he has stolen her prize zucchini.

Stalnaker's principal concern in connection with this example was to show how earlier belief attributions in a discourse can serve as contexts for the interpretation of attributions made in subsequent sentences. (He calls such contexts "secondary contexts", to distinguish them from the "primary context" of a given utterance, which contains information about the ways of the world that the utterance as a whole is about.) In the case of (220) the first sentence will enrich the primary context with the information that Phoebe believes that a man has broken into her garden. At the same time the sentence introduces a secondary context, viz. Phoebe's "current belief context", which is to the effect that a man broke into Phoebe's garden. The point of the secondary context is that it can serve the interpretation of cross-sentential devices (such as anaphoric pronouns) occurring in the complement sentences of following attitude attributions, in much the same way that primary contexts serve this purpose for occurrences of those devices when they occur outside the complements of attitudinal predicates.

Where there are two or more contexts to choose from, it is to be expected that the options for presupposition justification increase. And indeed they do. However, the addition of interpretational possibilities is much more dramatic than the availability of several justification contexts might have

[90] There has been a tendency in the philosophical and also in the linguistic literature to restrict the discussion of attitude attributions to sentences of this kind. But the repertoire natural languages make available for such purposes is much richer, including nouns such as *rumour*, *thought*, *opinion* or *fact*, adjectives such as *suspected* or *alleged*, prepositions such as *according to*. It is true that for many of the basic issues which attitude attributions raise the exclusive focussing on verbs is not a problem. But from a linguistic perspective such a narrow focus seems neverhteless arteficial and provincial.

suggested by itself. In particular, both pronouns and definite descriptions occurring in the complements of attitudinal verbs and *verba dicendi* come with a much wider repertoire of possible interpretation strategies than they do when they do not occur within the scope of such verbs. (Actually this is a more general phenomenon, which holds for a much wider range of expressions than just pronouns and descriptions, but we will explore it here only in connection with these.) To our knowledge the details of this problem have not been very systematically investigated, and our own observations here are of an exploratory character. Nevertheless they will keep us occupied for some time. Even the discussion of the seemingly simple (220), with which we begin our exploration and which illustrates only some of the issues that will preoccupy us in this final part of the present section, will take longer than might have been expected.

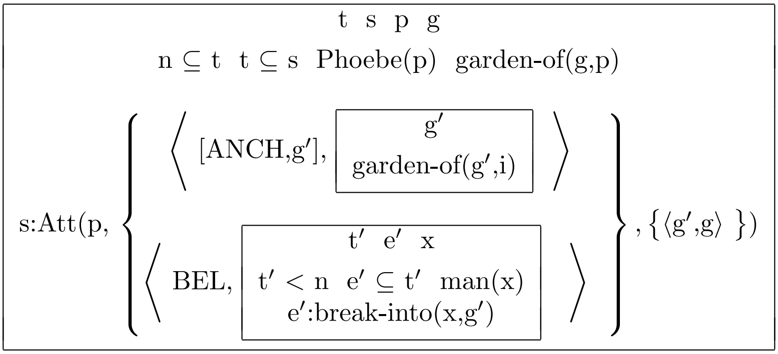
In the representation format we have developed in this chapter secondary contexts are identifiable as the second arguments of the DRT predicate Att. To see what this comes to in the case of (220) let us assume without further argument that its first sentence gets the representation given in (221).[91]

---

[91]We have made the plausible assumption that Phoebe has an internally and externally anchored representation for her garden. The discourse referent g serves as external representation of the object to which her representation g′ of her garden is anchored. We have also assumed that it is part of the internal anchoring information connected with this internal anchor that the object represented by g′ is "understood" by Phoebe as her garden. What other information the anchor contains — e.g. whether it is perceptual, based on memory or whatever — our representation leaves open; nothing in (220) indicates what this information might be like and there is no need for it to be made explicit. (In fact, with familiar objects, such as your own garden, your cat, your lover, your bed, etc. the notion of an anchor needs further scrutiny. Such objects do not have a single anchor, but an indefinite bunch of them, with each new contact between agent and object extending the bunch with a further component. It might be held that after only a little while the anchors within such a bunch blend into a single "super anchor". For current purposes such super anchors play the same role as anchors based on single encounters, so we refrain from pursuing the differences.) Note that the possession relation between Phoebe's garden and Phoebe is represented differently in relation to g′ and in relation to g. The external description of the object represented by g as Phoebe's garden makes use of the external representation p for Phoebe, whereas the internal representation of this information, as part of the internal anchor for Phoebe's own representation of her garden makes use of the indexical discourse referent i. We will return to this point when discussing the interpretation of the description "her prize zucchini", which is part of the second sentence of (220).

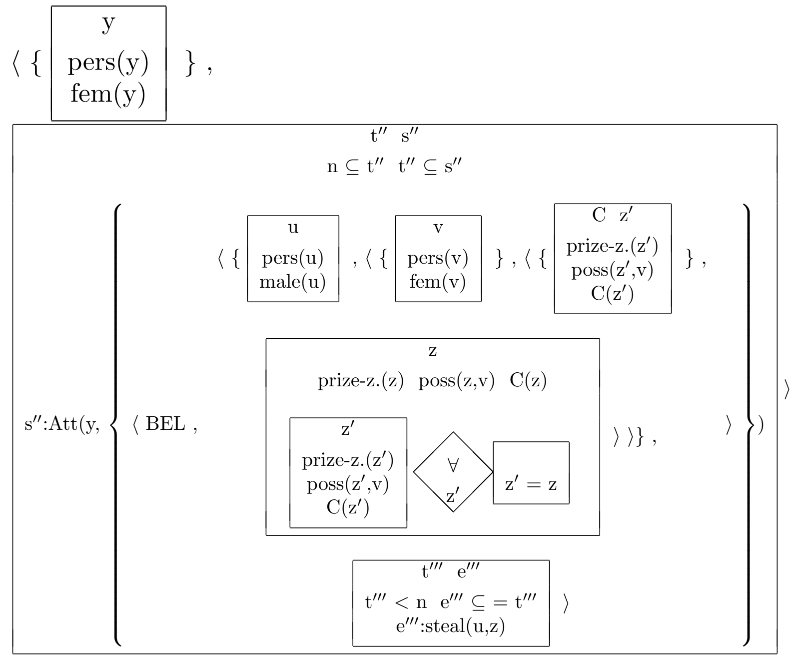
A further question that can be asked in connection with (221) is this: should we assume that Phoebe's representation of what the report describes with the NP *a man* is anchored too, either just internally or else both internally and externally? The case we are thinking of is one in which the reported beliefs of Phoebe's are a figment of her imagination, and that there is no particular man to whom her entity representation x can be seen as externally anchored. This still leaves open two possibilities: (i) that x is internally (though not externally) anchored; (ii) that x is not anchored at all. In (221) we have assumed that there is neither an external nor an internal anchor, but for the point that the example is meant to illustrate here it is not important how the question is settled. It would be of importance if we assumed that only anchored discourse referents can serve as the antecedents of subsequent pronouns. [Rooy, 1997].

(221)



The preliminary representation of the second sentence of (220) is given in (222)

(222)



This preliminary representation has five presuppositions. The first of these, triggered by the pronoun *she*, is adjoined to the DRS for the entire sentence. In addition, there are four other presuppositions, which are adjoined to the representation of the complement clause of *thinks*. Three of these are triggered by NPs: the pronoun *he*, the definite description *her prize zucchini* and the pronoun *her* inside it; the fourth is the presupposition on the contextual restrictor C for the existence- and-uniqueness condition from the presupposition triggered by the definite description (See Section 4). The presuppositions triggered by *her* and C are subordinate to the one triggered by *her prize zucchini.*

she. The presupposition for *she* can be resolved in the "primary" context provided by (221). (The secondary context is not available in this case, see below.) This means that only the discourse referents in the main universe of (221) are potential antecedents for the discourse referent y representing *she.* Resolution follows the pattern described in Section 4.2 and needs no further comment.

Resolution of the presuppositions adjoined to the representation of the complement is possible in principle both with respect to the primary and to the secondary context. We will take these possibilities in turn. But first we must address a general question concerning the role of the secondary context in presupposition resolution. This point will also be important in connection with the next two examples, which will be discussed in the following two subsections.

In order that the representation K' of the contents of a mental state can serve as interpretation context for the complement of an attitude attributing sentence S, it must be possible to see the attitude as a further component of the mental state which K' (partially) represents. This entails (i) that the agent to whom the attitude is attributed is the same as the agent of the mental state, and (ii) that the attitude is attributed to the agent at a time when he is in the mental state rerpesented by$K$': mental state and attitude must be simultaneous. In the case at hand the mental state is given as consisting of just one belief, represented by the DRS K' which occurs as second component in the second argument $\{\langle \mathrm{BEL}, K' \rangle\}$ of Att in (221). That the agent of the attribution made by the second sentence of (220) is the same as the agent of this belief follows when y is resolved to p. Simultaneity of attribution and context belief rests on the fact that both sentences of (220) are in the present tense. Thus both t and t'' must include the speech time n of (220). (As it stands, this doesn't strictly speaking entail that $t'' = t$. What it does entail is that there exists a time t'' during which the attribution made by the second sentence holds, which includes n and is included in t. But that seems enough to capture the content of (220).[92])

Having identified y with p and t' with t we have made the DRS of the complement clause of (221) available for resolution of the remaining presuppositions of (222). But this doesn't *by itself* answer the question how these presuppositions are to be resolved. This is true in particular of the three

[92]The possibility of identifying t' with t would be a consequence of treating present tense sentences as having an anaphoric dimension: apart from the requirement that the location time of the described eventuality include n, such a treatment would create the possibility of identifying this location time with some other time t which also includes n and which has already been introduced into the context. We will not elaborate this treatment further here. For the anaphoric dimension of tense see Section (3.5).

remaining NP presuppositions. In fact, we will see that each of these raises its own problems.

he. The least problematic is the anaphoric presupposition triggered by *he*. Now that we have secured the belief representation in (221) as (secondary) context for the interpretation of the complement of the second sentence of (220), the discourse referents in the universe of that representation are available as possible antecedents for the discourse referent u representing *he*. The obvious choice is x. So we resolve the presupposition by identifying u with x. (N.B this solves the problem with which the paper of Stalnaker in [Grimm and Merrill, 1988] and the comments by Kamp in that volume were principally concerned.)

her. Next, we turn to the possessive pronoun *her* of the definite description *her prize zucchini*. Intuitively it seems clear that this presupposition should be resolved by identifying the discourse referent v which represents *her* with the discourse referent p of the primary context. But such identifications come with a complication. By identifying the discourse referent v with one that is bound inside the main DRS we turn the propositional content of the attributed attitude into a singular proposition. In fact, the link between the discourse referent originating within the representation of the attributed content and the one bound outside this representation can be seen as an external anchor for the former. According to the position adopted in Section **????** this is coherent only if the internal discourse referent is internally anchored as well as externally. Thus, if we stick to the principle that there can be no external anchor without an internal anchor, then we must assume that by resolving v through identification with an external discourse referent, the interpreter is committed to the assumption that there is an internal anchor for v.

In cases where the external discourse referent with which the internal discourse referent v is identified represents an individual distinct from the agent of the attitude to whose representation v is internal, what was said in the last paragraph is all that needs to be said. But the situation has an additional complexity when the external discourse referent represents the agent. In this case it is also possible to interpret the internal discourse referent through identification with the external one. In the example we are discussing this amounts to identifying v with p — or, more fully, to taking v to be internally anchored and externally anchored to p. Note however that while the result of this is a representation of a belief of Phoebe's that is *de re* with respect to Phoebe herself, it is only one of two ways in which the pronoun *her* can be interpreted as referring to Phoebe.

The other possibility is to interpret *her* as signalling reference to Phoebe's self from her own internal perspective. We discussed self-reference in thought of this type in Section 5.3 and there we decided to represent it with the help of the special indexical discourse referent i. In line with that decision, the interpretation of which we are speaking now should involve identification of the discourse referent v with i. When a pronoun (or its representing discourse referent) is interpreted in this way, however, then there is of course no need for further assumptions about internal or external anchors. (For any occurrence of i is itself internal and, because of its direct link with the agent, it can be considered to have both an internal and an "external" anchor no matter what.) Below we will display both the *de se* and the *de re* interpretation of *her*.

her prize zucchini. The difference between the *de re* and the *de se* interpretation of *her* has its repercussions for that of the NP *her prize zucchini* which contains it. Let us begin with the assumption that *her* is given a *de re* interpretation. In that case there are still two options for the justification of the existence-and-uniqueness presupposition: either at the level of the secondary context or at that of the primary one. The first option amounts to assuming (i.e. accommodating the assumption) that Phoebe takes it that there is a unique entity x which is a garden, satisfies some additional predicate C and stands in the relation of being "had" by the person v of whom she has some anchored representation (which happens to be Phoebe). The second option amounts to there being a unique entity that is a garden, satisfies C and stands in the "being had" relation to the individual represented by p, i.e. to Phoebe. A further effect of this second option is that the discourse referent z which represents the denotation of the description moves to a position that is external to the representation of the belief. Once again this entails, in the light of our earlier assumptions about external and internal anchors, that there must be an internally anchored discourse referent that stands within the belief representation and that is externally anchored to z.

C. Let us, before we go on, display the representations to which these interpretational decisions lead. To do so, we also need to make a decision about the interpretation of C, but in connection with the example before us this is a matter that can be dealt with straightforwardly: the predicate "is a's garden" is uniquely satisfied for many values of a. If we are prepared to suppose that this is the case in particular for Phoebe, then the default interpretation of C as the universal predicate will serve. Let us assume that this is the way in which C in (221) gets resolved. Since this resolution makes the predications involving C vacuous, they will be dropped in the final representations of (220) that will be displayed below.

Both (223) and (224) assume the *de re* interpretation of *her*. (223) gives the internal accommodation of the existence-and-uniqueness presupposition of *her prize zucchini*, (224) its external accommodation. Both representations give the merge of the new representation with (221). For easier reading we have kept the two components of the merge graphically separate.

(223)

$$\boxed{\begin{array}{c} \text{t s p g} \\ \text{n} \subseteq \text{t} \quad \text{t} \subseteq \text{s} \quad \text{Phoebe(p)} \quad \text{garden-of(g,p)} \\ \text{s:Att(p,} \left\{ \begin{array}{c} \left\langle \text{[ANCH,g}'\text{],} \boxed{\begin{array}{c} \text{g}' \\ \text{garden-of(g}'\text{,i)} \end{array}} \right\rangle \\ \left\langle \text{BEL,} \boxed{\begin{array}{c} \text{t}' \quad \text{e}' \quad \text{x} \\ \text{t}' < \text{n} \quad \text{e}' \subseteq \text{t}' \quad \text{man(x)} \\ \text{e}'\text{:break-into(x,g}'\text{)} \end{array}} \right\rangle \end{array} \right\} \text{, } \{\langle \text{g}',\text{g} \rangle\} \text{)} \end{array}}$$

$$\boxed{\begin{array}{c} \text{t}'' \quad \text{s}'' \quad \text{y} \quad \text{v} \\ \text{n} \subseteq \text{t}'' \subseteq \text{s}'' \quad \text{y = p} \quad \text{v = p} \quad \text{t}'' = \text{t} \\ \text{s}''\text{:Att(y,} \left\{ \begin{array}{c} \left\langle \text{[ANCH,v}'\text{],} \boxed{\begin{array}{c} \text{v}' \\ \text{pers(v}'\text{)} \\ \text{fem(v}'\text{)} \end{array}} \right\rangle \\ \left\langle \text{BEL,} \boxed{\begin{array}{c} \text{z}' \\ \text{prize-z.(z}'\text{)} \quad \text{poss(z}'\text{,i)} \\ \boxed{\begin{array}{c} \text{z}'' \\ \text{prize-z.(z}''\text{)} \\ \text{poss(z}''\text{,i)} \end{array}} \; \begin{array}{c} \forall \\ \text{z}'' \end{array} \; \boxed{\text{z}'' = \text{z}} \end{array}} \right\rangle \\ \left\langle \text{BEL,} \boxed{\begin{array}{c} \text{t}''' \quad \text{e}''' \quad \text{u} \\ \text{t}''' < \text{n} \quad \text{e}''' \subseteq \text{t}''' \quad \text{u = x} \\ \text{e}'''\text{:steal(u,z}'\text{)} \end{array}} \right\rangle \end{array} \right\} \text{, } \{\langle \text{v}',\text{v} \rangle\} \text{)} \end{array}}$$

$$
\boxed{\begin{array}{c}
\text{t s p g}\\
\text{n} \subseteq \text{t}\quad \text{t} \subseteq \text{s}\quad \text{Phoebe(p)}\quad \text{garden-of(g,p)}\\[1em]
\text{s:Att(p,} \left\{\begin{array}{c}
\left\langle \text{[ANCH,g}'\text{]}, \boxed{\begin{array}{c} \text{g}' \\ \text{garden-of(g}'\text{,i)} \end{array}} \right\rangle \\[1em]
\left\langle \text{BEL}, \boxed{\begin{array}{c} \text{t}'\ \ \text{e}'\ \ \text{x} \\ \text{t}' < \text{n}\quad \text{e}' \subseteq \text{t}'\quad \text{man(x)} \\ \text{e}'\text{:break-into(x,g}'\text{)} \end{array}} \right\rangle
\end{array}\right\}, \{\langle \text{g}',\text{g}\rangle\})
\end{array}}
$$

(224)
$$
\boxed{\begin{array}{c}
\text{t}''\ \ \text{s}''\ \ \text{y v z}\\
\text{n} \subseteq \ \text{t}'' \subseteq \text{s}''\quad \text{y} = \text{p}\quad \text{v} = \text{p}\quad \text{t}'' = \text{t}\quad \text{prizez.(z)}\quad \text{poss(z,v)}\\[1em]
\boxed{\begin{array}{c} \text{z}'' \\ \text{prize-z.(z}''\text{)} \\ \text{poss(z}''\text{,v)} \end{array}} \left\langle \begin{array}{c} \forall \\ \text{z}'' \end{array} \right\rangle \boxed{\text{z}'' = \text{z}}\\[1em]
\text{s}''\text{:Att(y,} \left\{\begin{array}{c}
\left\langle \text{[ANCH,z}'\text{]}, \boxed{\begin{array}{c} \text{z}' \\ \ \end{array}} \right\rangle \\[1em]
\left\langle \text{BEL}, \boxed{\begin{array}{c} \text{t}'''\ \ \text{e}'''\ \ \text{u} \\ \text{t}''' < \text{n}\quad \text{e}''' \subseteq \text{t}'''\quad \text{u} = \text{x} \\ \text{e}'''\text{:steal(u,z}'\text{)} \end{array}} \right\rangle
\end{array}\right\}, \{\langle \text{z}',\text{z} \rangle\})
\end{array}}
$$

Note that in (224) there is no internal anchor for the referent of the pronoun *her*. The "external" interpretation of the definite description *her prize zucchini* which (224) represents allows for a purely external representation of *her*; only the discourse referent z for the entity denoted by the expression as a whole enters (indirectly via z′) into the content representation of the belief that the second sentence attributes.

The official notation in which these last two representations are given has the merit of making the distinction between internal and external anchors explicit. But it is cumbersome, and now that we have repeatedly demonstrated how it works, the time is ripe for simplifying it. We simplify by adopting the very notation against which we warned above: the one in which a discourse referent which is bound in a position external to
Att" occurs as an argument in one or more DRS-conditions which are within the scope of this occurrence. The use of this notation is now to be understood, however, as shorthand for the more complex one which appears in (223) and (224):

NOTATIONAL CONVENTION 70. Externally Bound Discourse Referents

If a condition P(w) is part of the representation K of an attitude content and w is bound outside the Att-condition of which K is an immediate constituent — i.e. the condition is of the form “s: Att(a,K,EA)”, while $\langle$ MOD,K $\rangle \in K$ for some MOD — then this is to be understood as equivalent to the condition “s: Att(a, $K \cup \{\langle$ [ANCH,w′], $K_{w'} \rangle \}$, EA $\cup \langle$ w′,w $\rangle$)”, where w′ is a new discourse referent (i.e. one that does not occur in the representation of which the Att-condition is part) and $K_{w'}$ is the DRS $\langle \{w'\}, \emptyset \rangle$.

In this simplified notation, the lower part of (223) (corresponding to the contribution of the second sentence of (220)) takes the form (225)

$$
(225)\quad
\boxed{\begin{array}{c}
t'' \;\; s'' \;\; y \;\; v \\
n \subseteq t'' \subseteq s'' \quad y = p \quad v = p \quad t'' = t \\
s'':\mathrm{Att}(y, \left\{ \begin{array}{c}
\left\langle \mathrm{BEL}, \boxed{\begin{array}{c} z \\ \text{prize-z.}(z) \;\; \mathrm{poss}(z,v) \\ \boxed{\begin{array}{c} z' \\ \text{prize-z.}(z') \\ \mathrm{poss}(z',v)\end{array}} \; \underset{z'}{\forall} \; \boxed{z' = z} \end{array}} \right\rangle \\
\left\langle \mathrm{BEL}, \boxed{\begin{array}{c} t''' \;\; e''' \;\; u \\ t''' < n \;\; e''' \subseteq t''' \;\; u = x \\ e''':\mathrm{steal}(u,z) \end{array}} \right\rangle
\end{array} \right\})
\end{array}}
$$

When *her* is interpreted *de se*, then of the two possibilities of the last paragraph for justifying the existence-and-uniqueness presupposition of *her prize zucchini* only the first one is a formal option. For the discourse referent i that is used to interpret the pronoun is internal to the representation; if we export the existence-and-uniqueness condition to the level of the primary context, then the condition “v = i” would have to be left behind and v would no longer be properly bound. The new (lower) part of the representation is given in (226)

$$
(226)\quad
\begin{array}{|c|}\hline
t'' \quad s'' \quad y \\
n \subseteq t'' \subseteq s'' \quad y = p \quad t'' = t \\
s'':\mathrm{Att}(y, \left\{ \begin{array}{c}
\left\langle \mathrm{BEL}, \begin{array}{|c|}\hline
z \\
\text{prize-z.}(z) \quad \mathrm{poss}(z,i) \\
\begin{array}{|c|}\hline z' \\ \text{prize-z.}(z') \\ \mathrm{poss}(z',i) \\ \hline \end{array}
\left\langle \begin{array}{c} \forall \\ z' \end{array} \right\rangle
\begin{array}{|c|}\hline z' = z \\ \hline \end{array} \\ \hline
\end{array} \right\rangle \\
\left\langle \mathrm{BEL}, \begin{array}{|c|}\hline
t''' \quad e''' \quad u \\
t''' < n \quad e''' \subseteq t''' \quad u = x \\
e''':\mathrm{steal}(u,z) \\ \hline
\end{array} \right\rangle
\end{array} \right\}) \\ \hline
\end{array}
$$

N.B We have adopted in this representation the convention of replacing the discourse referent v for *her* everywhere by i rather than link it to i via the condition "v= i".

The representations (223) — (226) may seem to present us with a problem in that none of them seem to fully capture the interpretation which a normal speaker is likely to understand the definite description of the second sentence of (220): as an NP that is to be interpreted *de re*, while the pronoun *her* that it contains is given a *de se* interpretation. None of the representations we have given shows this combination.

This doesn't mean that a representation with these properties would be incompatible with the principles we have formulated so far. For one thing, the representations which we have shown are to be understood as minimal, in the sense that the information they contain must obtain if the represented sentence (on its given interpretation) is to be true. They do not exclude the possibility that the interpreter's representation gets further enriched on the strength of various "pragmatic" considerations, which go beyond that which is conveyed by linguistic form as such. In the present case, however, it might even be argued that the interpretation that we are looking for is the result of yet another way of interpreting the definite description *her prize zucchini*, according to which it has a double function - — first as a means of identifying the external anchor of an internally anchored representation of the agent Phoebe, and secondly as a description of the information which she herself uses to represent the referent; it is in this second capacity that the description allows - and suggests — a *de se* interpretation of *her*, while it is the first function which makes the belief *de re* with respect to the zucchini. (227) gives the representation which results if the description is taken to play this double role. As (227) shows, the processing rule which reflects the double role interpretation of the NP must produce the effect that the internally anchored discourse referent introduced by the "internal" interpretation of the NP is externally anchored to the discourse referent established by its *de re* interpretation. In the simplified notation used in (227) this means that the conditions yielded by the internal interpretation take the form of predications of the external anchor

(227)
$$\boxed{\begin{array}{c} t'' \quad s'' \quad y \quad v \quad z \\ n \subseteq \ t'' \subseteq s'' \quad y = p \quad v = p \quad t'' = t \quad \text{prize-z.}(z) \quad \text{poss}(z,v) \\ \boxed{\begin{array}{c} z' \\ \text{prize-z.}(z') \\ \text{poss}(z',i) \end{array}} \left\langle \begin{array}{c} \forall \\ z' \end{array} \right\rangle \boxed{z' = z} \\ s'':\text{Att}(y, \left\{ \left\langle \text{BEL}, \boxed{\begin{array}{c} t''' \quad e''' \quad u \\ t''' < n \quad e''' \subseteq t''' \quad u = x \\ e''':\text{steal}(u,z) \quad \text{prize-z.}(z) \quad \text{poss}(z,i) \\ \boxed{\begin{array}{c} z'' \\ \text{prize-z.}(z'') \\ \text{poss}(z'',v) \end{array}} \left\langle \begin{array}{c} \forall \\ z'' \end{array} \right\rangle \boxed{z'' = z} \end{array}} \right\rangle \right\}) \end{array}}$$

Summary of 5.4: Our discussion of example (220) has focused on two aspects of the interpretation of attitude attribution sentences. First — this is a very general point, which will play a major role in the examples considered in the next two sections — in order that representations of mental states in the context in which such a report is interpreted can serve as "secondary contexts", it must be established that they represent mental states of the agent to whom the report attributes a state and moreover that they hold at the time at which the agent has this attitude according to the report.

The second point, to which most of the discussion of this section has been devoted, concerned the multiplicity of possible interpretations for certain NPs which arise when the NP is part of the complement clause of an attitudinal predicate (such as, in our example, the verbs *believe* and *think*). Although we haven't explicitly stated the interpretation (= DRS construction) rules for NPs which cover these new possibilities, we trust that the discussion has given a fairly clear indication how rules could be stated which lead to the representations we have shown. Now that we have seen in some detail to what multiplicity of alternative interpretation rules pronouns and definite descriptions give rise when they occur in the scope of attitude predicates, it is well to reiterate our earlier observation that this dramatic increase in intepretational options is by no means limited to just these two types of expressions. WE find a comparable increase for other NP types that have anaphoric uses (such as demonstrative NPs), as well as —

and this is particularly important — for indefinite NPs. (It is an old observation about indefinite NPs in the complements of attitude predicates like "believe" that they usually allow for a "*de re*" as well as a"*de dicto*" interpretation. The *de re* option can be seen as one way in which indefinite NPs can be "specific".) Moreover, new interpretational distinctions also arise for types of expressions other than NPs. for these reasons extending the construction algorithm for a language fragment without attitude predicates to one which includes them is a complicated matter, which requires careful analyses of what ranges of representations are possible for which sentences.

In the next two sections we consider examples which illustrate two further aspects of the representational capacities of our formalism and of the interpretational principles needed to interpret sentences and discourses which make use of these possibilities. The example of the next section concerns the description of attitudinal change, i.e. of a temporal succession of distinct mental states of the same agent. The section after that is devoted to a case in which attitudes are attributed to two different agents who can be assumed to share a certain common ground. Such a common ground will often make it possible to use an attribution that has been made to one of them as context for the interpretation of an attribution that is made subsequently to the other.

Reporting Changes of Attitudinal States

The next example illustrates the ability of the present formalism to accurately represent temporal relations between the times at which attitudes are entertained and the times of the eventualities mentioned in the propositional contents of those attitudes.

(228) On Sunday Bill heard that Mary was in Paris. On Tuesday he learned that on the previous day she had left.

(228) also exemplifies some of the complexities that arise in connection with the interpretation of tenses and other expressions referring to time; these, we will see, have much to do with the way in which the temporal aspects of the contents of thoughts are connected with the times at which they are entertained.

The first instance of this problem that we must consider here is the past tense in the complement of *heard* in the first sentence of (228). It is a well-known and much discussed fact of languages like English that a simple past tense within the complement of an attitudinal verb which itself is also in a past tense can be understood as expressing simultaneity between the eventuality of the complement and the attitude or attitudinal change referred to by the verb itself — a phenomenon known in the literature as "sequence of tense". This is not the only possible interpretation for the past tenses of verbs in the complements of past tense attitude verbs or verba dicendi; they can also be understood as expressing anteriority to the time of the matrix verb eventuality. Thus the first sentence of (228) can be understood not only as saying that what Bill came to believe on Sunday was that Mary was in Paris at the very time he had just formed this new belief, but also that what he came to believe was that she was in Paris at some time before that when his new belief came about. How the tense of the complement is interpreted will depend on several factors, one of which is the Aktionsart of the embedded verb. If this is an event verb, then in English the simultaneous (i. e. sequence of tense) interpretation is excluded. (For the same reason that the use of the simple present is proscribed in normal context. However, the simultaneous reading returns as a possibility when the simple past is replaced by a past progressive, just as present progressives of event verbs are acceptable in normal contexts.) When the embedded verb is stative, then its simple past will in general be ambiguous between the simultaneous interpretation and the anterior interpretation. (When the verb phrase of the complement sentence is stative, there is usually a preference for the simultaneous interpretation.) We choose the simultaneous interpretation for the first sentence as the basis for the interpretation of the second sentence, which is the real topic of this subsection. (See (229), (230) below).

How do we represent the simultaneous interpretation? For the most part this should be clear from what has been said about the representation of propositional attitudes so far. In particular, simultaneity of the content of a thought with the time when the thought is being entertained can be expressed with the help of the temporal indexical n. In other words, the state of Mary's being in Paris is to be represented as surrounding the "internal present" which is denoted by an occurrence of n within the representation of Bill's attitudinal state. As the sentence makes clear, this attitudinal state is temporally located within the interval denoted by *Sunday.* (See Section (3.5) for details.)

New in the representational challenge which the first sentence of (228) presents is the representation of the verb *hear.* Like the verb *learn* of the second sentence, *hear*, in the use that is made of it in (228), conveys the emergence of a new belief (or item of knowledge), where just before there wasn't such a belief (or perhaps even a contrary opinion). It is arguable that this bit of information, which is unequivocally part of the meaning of *learn*, is only an implicature in the case of *hear.* But we will leave this question — whether we are dealing with an implicature or a genuine part of the lexical meaning — for some other occasion, and assume for simplicity that there is a lexical meaning of *hear* which does include the previous non-existence of the belief as a component, just as this is the case for *learn.* (We also ignore that *hear*, as opposed to *learn*, carries implications about the way in which the information reaches the agent — for instance, hearing is not the same as (learning by) reading.) We also pass over the question whether "x heard that p" really entails that x came to believe that p. Perhaps Bill heard that

Mary was in Paris, but didn't believe a word of it? In the context of (228), where the second sentence seems to refer back to the attitudinal state which has been set up by the first, this second possibility seems more remote than it may be in other contexts, and so it too is set aside.

The relevant reading of *hear*, then, which we assume to be the one relevant to the present sentence is that of a change-of-state verb, which expresses a transition from the state of not believing/knowing the content of what one hears to the state in which one does believe/know that. We also make the usual assumption about pre-states of change-of-state verbs, viz. that such verbs carry a presupposition to the effect that a pre-state of the relevant type (one which denies the type of the result state) obtains at the time when the venetuality described by the verb begins.[93]

We are now ready to present the representation of the first sentence of (228). (229) gives the preliminary representation for this sentence, with explicit and separate representations of the presuppositions triggered by the proper names *Sunday, Bill, Mary* and *Paris* and the pre-state presupposition of *hear*. In the final representation (230) for the sentence all five presuppositions have been accommodated.

[93]There is one further issue connected with the first sentence of (228) which we must briefly comment on before showing the representation which we will use as context for the interpretation of the second sentence. This issue concerns the representation of the names *Mary* and *Paris*. It doesn't have to do with the temporal aspects of (228) as such and we would have raised it in connection with our last example if that had happened to contain a proper name within an attitudinal complement. Occurrences of names within the complements of attitude verbs and dicendi verbs are typically understood as de re. (There may be marginal exceptions to this, but if we are right, then these really are marginal.) This means that the reported belief must be construed as involving discourse referents which are externally anchored to the person Mary and the city of Paris, respectively.

We assume that the same is true for the NP *Sunday*. Weekday names aren't proper names in the sense of having all properties that semanticists and philosophers of language take to be part of the concept of a proper name. In particular, the denotations of weekday names depend in systematic ways on the contexts in which they are used. We will ignore this contextual dimension of the reference of *Sunday* here. (No information about the context was given anyway. We will take *Sunday* in (228) to refer the last Sunday before the utterance time, but nothing much hangs on this.) What is more relevant to what will be said about the interpretation of (228) below is the temporal relation between the referents of *Sunday* in the first sentence and *Tuesday* in the second. We will assume that *Tuesday* refers to the Tuesday immediately after the referent of *Sunday* in the first sentence.

(229)

$$\left\langle \left\{ \boxed{\begin{array}{c} t_0 \\ \text{Sunday}(t_0) \end{array}}\ ,\ \boxed{\begin{array}{c} b \\ \text{Bill}(b) \end{array}}\ ,\ \boxed{\begin{array}{c} m \\ \text{Mary}(m) \end{array}}\ ,\ \boxed{\begin{array}{c} p \\ \text{Paris}(p) \end{array}} \right\}, \right.$$

$$\left\langle \left\{ \boxed{\begin{array}{c} s_0 \\ s_0 : \neg \boxed{\begin{array}{c} s_0' \\ s_0' \subseteq s_0 \\ s_0'\text{:Att(b, } \left\{ \left\langle \text{BEL}, \boxed{\begin{array}{c} s' \\ n \subseteq s' \\ s'\text{:IN(m,p)} \end{array}} \right\rangle \right\} \text{)} \end{array}} \end{array}} \right\}, \right.$$

$$\left. \boxed{\begin{array}{c} t_1 \quad e \quad s_1 \\ t_1 \subseteq t_0 \quad e \subseteq t_1 \quad s_0 \supset\!\subset e \supset\!\subset s_1 \\ s_1\text{:Att(b, } \left\{ \left\langle \text{BEL}, \boxed{\begin{array}{c} s' \\ n \subseteq s' \\ s'\text{:IN(m,p)} \end{array}} \right\rangle \right\} \text{)} \end{array}} \right\rangle \right\rangle$$

(230)

$$\boxed{\begin{array}{c} t_0 \quad b \quad m \quad p \quad s_0 \quad t_1 \quad e \quad s_1 \\ \text{Sunday}(t_0) \quad \text{Bill}(b) \quad \text{Mary}(m) \quad \text{Paris}(p) \quad t_1 \subseteq t_0 \quad e \subseteq t_1 \quad s_0 \supset\!\subset e \supset\!\subset s_1 \\ s_0 : \neg \boxed{\begin{array}{c} s_0'' \\ s_0'' \subseteq s_0 \\ s_0'\text{:Att(b, } \left\{ \left\langle \text{BEL}, \boxed{\begin{array}{c} s' \\ n \subseteq s' \\ s'\text{:IN(m,p)} \end{array}} \right\rangle \right\} \text{)} \end{array}} \\ s_1\text{:Att(b, } \left\{ \left\langle \text{BEL}, \boxed{\begin{array}{c} s' \\ n \subseteq s' \\ s'\text{:IN(m,p)} \end{array}} \right\rangle \right\} \text{)} \end{array}}$$

(Here "$\supset\!\subset$" denotes abutment of two eventualities or the periods (see Section (3.5)). Note the somewhat cumbersome way of expressing the information that Bill's belief that Mary is in Paris is new: the event e contributed by *hear* is represented as the transition to a post-state s_1 in which Bill has a belief to the effect Mary is in Paris from the pre-state s_0 in which Bill's attitudes do not include such a belief. We will discuss the representation of state transitions at length in Section 6.

We now pass to the central concern of this subsection: the interpretation of the second sentence of (228) in the light of the context established by the first sentence (230). We split the discussion of the issues which need addressing into two parts, (i) the conditions that must be satisfied in order that the secondary context provided by the belief attribution of the first sentence can be used in the interpretation of the second sentence, and (ii) some of the complexities that arise in connection with the interpretation of certain constituents of the complement of the second sentence, given that both the primary and the secondary context are available for the resolution of presuppositions.

Temporal Alignment of the Secondary Context with the Attitude Report. The second sentence of (228) is in many ways like the first. But there is one crucial difference, and this is our principal reason for making (228) the topic of a separate discussion: as in our previous example (220), interpreting the complement of the matrix verb of the second sentence — here the verb *learn* — requires as context the representation provided by the complement sentence of the first sentence. As we noted in connection with (220) using one attitude attribution as context for another presupposes that the two attitudes must be part of a single attitudinal state. This entails that we must be dealing (i) with a single attitude bearer, and (ii) with a single time at which both attitudes are entertained. In our first example (220) verifying that these conditions were fulfilled was straightforward. Here it is not.

Note that what the first sentence of (228) tells us is just that Bill acquired a certain belief on Sunday. We are not told whether he kept this belief — that Mary was in Paris during some period including the time when he heard that she was in Paris — until the time on Tuesday, when he is said to have learned that she "left". Yet we must assume that he did, for otherwise it is hard to make sense of the belief attribution that is made in the second sentence: the intransitive verb *leave* always involves, from a semantic point of view, an argument for the place from which the subject leaves, irrespective of whether this place is mentioned explicitly (in the form of a *from*-PP) or not. Moreover, when the place is not mentioned explicitly, there is always an implication that it can be reconstructed from context.

In the case of (228) the resolution of this instance of "implicit argument anaphora" is intuitively clear: it seems clear that the place of which Bill learns that Mary left from there was Paris. Since a discourse referent representing Paris is present in the primary context that is given by (230), this resolution does not require the secondary context. But *leave* also comes with a pre-state presupposition, viz., that its subject was in the place that she is said to have left. In principle, pre-state presuppositions are quite easily accommodated, but nevertheless the use of *leave* (like that of other change-of-state verbs) creates a definite presumption of the relevant pre-state being "already known" — that is, part of the context. In the case at hand this means that there is a presumption that the pre-state — that of Mary being in Paris — is "already known" to Bill. The context provided by (230) supports this presumption, in that the obtaining of this pre-state is the content of its secondary context. However, the secondary context of (230) can resolve the pre-state presupposition triggered by *leave* only when it is assumed that the belief which (230) attributes to Bill continues to be his belief until the time on Tuesday when he finds out about Mary's departure. That this is really so cannot be strictly inferred from (230), but must be accommodated. It is the kind of accommodation that comes easily, since it is in line with a general principle of discourse "persistence": states of affairs which the discourse claims to obtain at some given time will typically be assumed to persist unless the discourse provides explicit or implicit information to the effect that the state has come to an end.[94] Nevertheless, it is an accommodation of some kind.

The next question we must address is what exactly is being accommodated. This may seem a strange question, with an answer that is entirely obvious: we simply add a condition that the state s_1 of (230) still holds at the time of the event e' of Bill learning that Mary has left Paris. But there is a subtlety here. It is certainly true that the accommodation just described is one way of arriving at a coherent interpretation to the second sentence. But it is not the only one. The belief which (230) attributes to Bill on Sunday is that Mary is in Paris *on Sunday.* There are two ways in which this belief can persist as time goes on, either as the belief that Mary is in Paris in the sense of the "psychological present", i.e. at the time at which the belief is entertained, or else as the belief that Mary was in Paris on Sunday. The accommodation mentioned above is to the effect that Bill's belief persists as a belief "about the present". For at the later time on Tuesday to which the accommodation extends the belief, the discourse referent n inside the characterisation of its content refers to this time on Tuesday, and not to the earlier time on Sunday, when Bill heard that Mary was in Paris. The more modest accommodation of the belief that Mary was in Paris on Sunday requires that the belief content now be represented in a different way, not as a "present tense" but as a "past tense belief": the time of the state of Mary being in Paris must now be represented as one before (the embedded occurrence of) n, rather than simultaneous with n. This second accommodation leaves it open whether Bill believed on Tuesday that Mary

[94]This principle, also called "monotonicity" (see [Reyle and Rossdeutscher, 2001]) is reminiscent of the frame problem from AI. But the discourse effect tends to be even stronger, for it is a constraint on discourse coherence that the termination of such states must be conveyed, if this is what the speaker or author intends. So the very fact that the discourse says nothing about termination can be taken as a sign that the state is to be understood as persisting.

was still in Paris then, whereas the first accommodation claims this. We will show both accommodations below. As we will see, they have slightly different consequences for the remaining aspects of the interpretation of the second sentence of (228).

A further observation concerns the use of the past perfect. It was observed in Section (3.5) that this tense is typically interpreted as involving a past Temporal Perspective Point, locating the described eventuality in the past of this TP-point. In the present case there are two possible choices for this TP-point, (i) the time of the event e introduced by *hear* in the first sentence and (ii) the event e′ introduced by *learn* in the matrix clause of the second sentence. The first choice places the event of Mary's leaving before Sunday. So, on this interpretation the information which Bill gets on Tuesday contradicts what he heard on Sunday. In view of the "corrective" character which (228) takes on with this interpretation, one would, if this had been the intended interpretation, have expected some kind of contrastive element, such as e.g. *but* as first word of the second sentence, to bear witness to the contrast between the claim made by the first sentence and that made by the second. So, without dwelling further on the general principle at work here, we take the absence of such a particle as a justification for choosing the second option, according to which the event of Mary's leaving Paris occurred before Tuesday.

A similar ambiguity arises also in connection with the interpretation of the temporal adverbial *the day before*. This adverbial has the form of a definite description, and its referent has to be determined accordingly. The descriptive content of this description is the relational expression *day before*. Like the verb *leave*, this phrase can occur either with an explicit second argument, as in *day before Sunday*, say, or without any phrase designating this argument. The latter possibility is the one we find realised in (228). And like with the verb *leave* there is in such cases an implication to the effect that the missing argument should be recoverable from the context. Moreover, when the phrase *the day before* occurs as adjunct to a finite VP, it is subject to a default recovery principle according to which the missing argument is the TP-point that is also needed to interpret the tense of te clause. This means that if we take the time of the event e as TP-point, then the *the day before* gets an interpretation on which it denotes the Saturday before the mentioned Sunday; and when the time of e′ is taken as TP-point, then the phrase is understood as denoting the following Monday. Since we have already decided to identify the TP-point with e′, we are led to interpret the description as denoting the Monday.

(231) gives the preliminary representation for the second sentence of (228), and (232) and (233) the updates of the context DRS (230) with the two mentioned accommodations. After these diagrams we will first have to say a few more things about presupposition justification of (231) on the basis of, respectively, (232) and (233). Only after that we will give,in (234)

and (235) the representations of (228) which result when all presuppositions have been resolved and the representation of the new sentence has been merged with the context representation.

$$\left\langle \left\{ \boxed{\begin{array}{c} t_0' \\ \text{Tuesday}(t_0') \end{array}}, \boxed{\begin{array}{c} u \\ \text{pers}(u) \\ \text{male}(u) \end{array}}, \boxed{\begin{array}{c} v \\ \text{pers}(v) \\ \text{fem}(v) \end{array}}, \boxed{\begin{array}{c} l \\ \text{loc}(l) \end{array}}, \boxed{\begin{array}{c} t_4 \\ \text{(impl. arg.)} \\ \text{“day before”} \end{array}} \right\}, \right.$$

(231)
$$\left\langle \left\{ \boxed{\begin{array}{c} s_2 \\ s_2 : \neg \boxed{\begin{array}{c} s_0' \\ s_2' \subseteq s_2 \\ \\ s_2'\text{:Att}(u,\{ \ \langle \text{BEL}, \langle \{ K_1 \}, K_2 \rangle \rangle \ \}) \end{array}} \end{array}} \right\}, \right.$$

$$\left. \boxed{\begin{array}{c} t_3 \ \ e' \ \ s_3 \\ t_3 \subseteq t_0' \ \ e' \subseteq t_3 \ \ s_2 \supset\!\subset e' \supset\!\subset s_3 \\ \\ s_3'\text{:Att}(u,\{ \ \langle \text{BEL}, \langle \{ K_1 \}, K_2 \rangle \rangle \ \}) \end{array}} \right\rangle \right\rangle$$

where K_1 is the DRS $\boxed{\begin{array}{c} s_3' \\ s_3'\text{:IN}(v',l') \end{array}}$

and K_2 is the DRS $\boxed{\begin{array}{c} t_0'' \ \ t'' \ \ e'' \ \ s_3'' \\ \text{day}(t_0'') \ \ t_0'' \subseteq \text{day}(t_4) \ \ t'' \subseteq t_0'' \\ t'' < n \ \ e'' \subseteq t'' \ \ s_3' \supset\!\subset e'' \supset\!\subset s_3'' \\ \\ s_3''\text{:}\neg\text{IN}(v,l) \end{array}}$

(232)
$$\boxed{\begin{array}{c} t_0 \ \ b \ \ m \ \ p \ \ s_0 \ \ t_1 \ \ e \ \ s_1 \\ \text{Sunday}(t_0) \ \ \text{Bill}(b) \ \ \text{Mary}(m) \ \ \text{Paris}(p) \ \ t_1 \subseteq t_0 \ \ e \subseteq t_1 \ \ s_0 \supset\!\subset e \supset\!\subset s_1 \\ \\ s_0 : \neg \boxed{\begin{array}{c} s_0' \\ s_0' \subseteq s_0 \\ \\ s_0'\text{:Att}(b, \left\{ \left\langle \text{BEL}, \boxed{\begin{array}{c} s' \\ n \subseteq s' \\ s'\text{:IN}(m,p) \end{array}} \right\rangle \right\}) \end{array}} \\ t_3 \subseteq s_1 \\ s_1\text{:Att}(b, \left\{ \left\langle \text{BEL}, \boxed{\begin{array}{c} s' \\ n \subseteq s' \\ s'\text{:IN}(m,p) \end{array}} \right\rangle \right\}) \end{array}}$$

$$(233)\quad (230)\quad \underline{\cup}\quad \begin{array}{|c|}\hline \mathrm{s}_4 \\ \mathrm{t}_3 \subseteq s_1 \\ \mathrm{s}_1{:}\mathrm{Att}(\mathrm{b}, \left\{ \left\langle \mathrm{BEL}, \begin{array}{|c|}\hline \mathrm{s}' \\ \mathrm{t}_0 < \mathrm{n} \\ \mathrm{t}_0 \subseteq \mathrm{s}' \\ \mathrm{s}'{:}\mathrm{IN}(\mathrm{m,p}) \\ \hline \end{array} \right\rangle \right\}) \\ \hline \end{array}$$

N.B. In both (232) and (233) the accommodation involves adding a condition which guarantees that the belief about Mary being in Paris lasts up to the time t_3 of the event e' of 231. In the case of (232) this can be represented simply by insisting that the very belief state s_1 of (230) overlaps with t_3. The case of (233) is somewhat more involved since here the representation of the content of the belief has to be modified so that it suits the new, later belief time t_3.

Presupposition Resolution for the Preliminary Representation of the Second Sentence of (228) As stated above, each of the updated contexts (232) and (233) makes it possible to justify the pre-state presupposition of *leave* in (231). (This requires that u be resolved to b, v to m, l to p and t_4 to t_1.) But there is an obvious difference between the two cases: the belief attributed to Bill in (233) is compatible with the belief that is attributed to him in (231), but the belief attribution of (232) is not. This means that the two interpretations corresponding to (232) and (233), while both possible, are conceptually quite different. If Bill was, at the time on Tuesday when he learned about Mary's departure, in the doxastic state indicated in (233), then it is reasonable to assume that his new doxastic state results from the immediately preceding one through simple addition of the new belief that Mary left Paris on Monday. If Bill's immediately preceding doxastic state is as described in (232), then addition to it of the new belief represented in (231) leads to a contradiction so obvious that it is hardly credible that Bill should have acquiesced in it. Almost certainly he will have revised his former beliefs in the light of what has just become known to him. The intuitively most likely revision would be that Mary didn't remain in Paris until Tuesday — in other words, that the state of her being in Paris did not persist as far into the future as Bill had erroneously supposed up to that point. This leads us back to (233), the result of the weaker accommodation of (230). After merging with the non-presuppositional part of (231), we get the representation given in (234).

$$(234)\quad (231) \uplus \boxed{\begin{array}{c} t_0' \quad u \quad v \quad l \quad t_4 \quad s_2 \\ \text{Tuesday}(t_0') \quad u = b \quad v = m \quad l = p \quad t_4 = s_2 \\ \\ s_2 : \neg \boxed{\begin{array}{c} s_2' \\ s_2' \subseteq s_2 \\ \\ s_2'\text{:Att}(u,\{ \ \langle \text{BEL}, \langle \{ K_1 \} , K_2 \rangle \rangle \ \}) \end{array}} \\ \\ \boxed{\begin{array}{c} t_3 \quad e' \quad s_3 \\ t_3 \subseteq t_0' \quad e' \subseteq t_3 \quad s_2 \supset\!\subset e \supset\!\subset s_3 \\ \\ s_2'\text{:Att}(u,\{ \ \langle \text{BEL}, \langle \{ K_1 \} , K_2 \rangle \rangle \ \}) \end{array}} \end{array}}$$

where again K_1 and K_2 are as under (231).

The representation in (234) seems very similar to that in (233). For one thing the two represent the same truth conditions. It should be stressed, however, that as *interpretations* of (228) they are clearly distinct. The stories that (232) and (233) tell about Bill up to the time when he learned that Mary left on the previous day differ in important details. That the representation in (234), which is based on the accommodation shown in (232), converges in the end with the one that is based on (227), depends crucially on the likely assumption that in the case of (232) Bill will have revised his earlier belief in the light of his new information. Belief revision, however, is something very different from what goes on when we arrive, by merely following the linguistic rules of interpretation, at a semantic representation that is inconsistent right away.

The discussion of this last section has demonstrated the same problems of exposition that became evident already in connection with hte last one: A large number of seemingly unrelated details, many of which also had no direct bearing on the issues which the example was meant to illustrate. We already drew attention to this at the outset of the last section; if we return to the observation once more here, it is in the hope that the reader is in a better position now to appreciate the quandary: Any example which illustrates the points on which the last two subsections were trained, will present a comparable range of issues, some closely related to the cantral issues and others hardly or not at all. But even those which are not or only distantly related require some attention if the representations proposed are to come across as well-motivated. Our discussions of the two examples of these last subsections would have made a much less haphazard impression, if it had been possible to rely on antecedently given solutions of all those problems which we encountered and which are irrelevant or ancillary to our principal cocerns. But this would have required a very different set-up of the present chapter, which in our own view would have made it quite unsuitable as a chapter for a Handbook. In the light of these considerations dealing with marginally relevant issues as they are tossed up by the examples chosen seemed to us the lesser of two evils.

Whether or not the mode in which we have proceeded in these sections is seen as satisfactory, there is an obvious moral that can be drawn: In order to provide a realistic account of the semantic representation of all but hte simplest sentences and discourses one needs to appeal to a highly complex system of interacting interpertation rules.

For a good number of the issues which we were forced to treat on the fly in dealing with our examples more systematic treatments can be found in the DRT literature than they could be given here. But this isn't the case for all of them. DRT may compare quite favourably with other frameworks for natural language semantics when it comes to coverage, but its coverage is still quite limited nonetheless. This means in particular that building a DRT-based semantics for a fragment of a natural language such as English which is large enough to permit relatively unimpeded use in all but the most special contexts remains a big challenge.

As regards the issues which have been our central concern in these last two subsections — viz. the ways in which successive attitude attributions can be semantically connected — there is a special reason why it is hard to come up with examples that illustrate the point without getting involved in additional problems. This is because so many natural examples which illustrate this kind of connectedness establish the connection by means of an anaphoric expression in the second attribution, which picks up the propositional content attributed by the first one. Some examples are given in 235.

(235) a. Bill thought that Mary was in Paris. But then he discovered that this wasn't so.

b. Bill thought that Mary was in Paris. But then he discovered that he was wrong.

c. For many years Bill wanted to make a trip to Egypt, but he doesn't want to any more now.

d. Bill very much wanted to prove that theorem and he was terribly pleased that he had when at last he had succeeded.

e. Bill had wanted to be a politician, but when he understood why he wanted this, his desire disappeared.

Pronominal and demonstrative reference to propositional attitudes and their contents is a subject in its own right, which goes well beyond what we have touched upon here. It is a topic that has received a considerable amount of attention in the DRT-related literature, cf. [Asher, 1993]. The same is true of ellipsis. Here too there is a growing literature (See [Hardt, 1992], [Asher *et al.*,], [Schiehlen, 1999]). (235.e), moreover, points up a problem of a different sort. Many of our attitudes are "second order" in that they are about some of our own attitudes. (A belief which you entertain about the origin of one of your desires is only one of a wide variety of different types of such second order attitudes.) It might be thought that such attitudes can be represented in much the same way in which the present formalism represents attitude attributions that one person makes to another (and about which we will have more to say in the next subsection). However, representing self-reflection in this way fails to capture one special feature of attributing properties to one's own thoughts. Such attributions have a kind of transparency that isn't there when we attribute thoughts to others. Whenever we attribute a thought to someone else, we must rely on hypotheses about what thoughts this person has. These hypotheses involve representations we form of the other's thoughts and to the question whether or how closely they capture the thoughts which we attribute to the other correctly there is rarely if ever a conclusive answer. But when we think about our own thoughts, then the subjects of our reflections are immediately accessible to us, with an immediacy that is reminiscent of how the direct access we have to our own selves. The thoughts that are formed in self-reflection are thus thoughts which are directly about the first order thoughts on which they are targeted; their contents are singular propositions whose subjects are other thoughts. But they are singular propositions of a special kind, similar to the singular proposition about my own self that is the content of the thought I have when thinking, say, "I want to go home".

Self-reflection is an important topic within the general theory of propositional attitudes and attitude attributions, but it is one we will not pursue here. A proposal for the representation of self-reflective thoughts within the present framework can be found in [Kamp, 1999].

Shared Attitudes between Different Agents

Our last example concerns the possibility of referentially connected attributions to different agents. Consider:

(236) Phoebe believes that a man broke into her garden and that he stole her prize zucchini. Ella thinks he didn't take anything.

The first sentence leads to the same representation as the two sentences of (220). One of these representations was given in (224) and we will assume that it is this representation which the interpreter of (236) assigns to its first sentence. (224) is repeated here as (237), in the abridged notation in which internal and external anchors are not explicitly mentioned and after merging of the representations of the two conjuncts of the first sentence into a single DRS.:

(237)
$$
\begin{array}{|c|}
\hline
\mathrm{t\ \ s\ \ p\ \ g\ \ \ y\ \ v\ \ z} \\
\mathrm{n} \subseteq \mathrm{t} \quad \mathrm{t} \subseteq \mathrm{s} \quad \text{Phoebe(p)} \quad \text{garden-of(g,p)} \\
\mathrm{v} = \mathrm{p} \quad \mathrm{t}'' = \mathrm{t} \quad \text{prize-z.(z)} \quad \text{poss(z,v)} \\
\\
\begin{array}{|c|}\hline \mathrm{z}'' \\ \text{prize-z.}(\mathrm{z}'') \\ \text{poss}(\mathrm{z}'',\mathrm{v}) \\ \hline \end{array}
\left\langle \begin{array}{c} \forall \\ \mathrm{z}'' \end{array} \right\rangle
\begin{array}{|c|}\hline \\ \mathrm{z}'' = \mathrm{z} \\ \hline \end{array} \\
\\
\text{s:Att(p,} \left\{ \begin{array}{l}
\left\langle \text{BEL,} \begin{array}{|c|}\hline \mathrm{t}'\ \ \mathrm{e}'\ \ \mathrm{x} \\ \mathrm{t}' < \mathrm{n} \quad \mathrm{e}' \subseteq \mathrm{t}' \quad \text{man(x)} \\ \mathrm{e}'\text{:break-into(x,g)} \\ \hline \end{array} \right\rangle \\
\\
\left\langle \text{BEL,} \begin{array}{|c|}\hline \mathrm{t}'''\ \ \mathrm{e}'''\ \ \mathrm{u} \\ \mathrm{t}''' < \mathrm{n} \quad \mathrm{e}''' \subseteq \mathrm{t}''' \quad \mathrm{u} = \mathrm{x} \\ \mathrm{e}'''\text{:steal(u,z)} \\ \hline \end{array} \right\rangle
\end{array} \right\}) \\
\hline
\end{array}
$$

(237) is the context of interpretation for the second sentence of (236). One of the questions which arise in connection with the interpretation of this sentence is on the face of it quite similar to a problem we encountered when discussing the second sentence of (228). There a persistence accommodation was necessary to extend the belief that the first sentence attributes to Bill at t_1 to the later time t_2, so that it could serve as context of interpretation for the belief attribution made by the second sentence. In (236) a similar problem arises in connection with the pronoun *he* in the second sentence. What does this pronoun refer to? “Well”, one might be inclined to reply, “to the man of whom Phoebe believes that he broke into her garden and stole her prize zucchini.” But how and in what sense can Ella’s thought be about this man, if, as we assumed in our discussion of (220), there is for all we know no such man in reality, and if what is said in the first sentence is a figment of Phoebe’s imagination? Clearly, the anaphoric relation between *a man* in the first and *he* in the second sentence of (236) makes no sense unless there is some mental content which Ella shares with Phoebe. What is needed, therefore, is an accommodation according to which some of what the first sentence attributes to Phoebe is also part of the beliefs of Ella.

But what exactly should be accommodated in this case? That is not so easy to say. On the one hand, as much should be accommodated as is necessary for a meaningful interpretation of *he*. On the other, the accommodation should be modest enough to avoid attributing to Ella beliefs that are so plainly contradictory that they could only be seen as incoherent. One possibility which meets these two conditions — but it is only one among several — is to accommodate the belief attributed to Phoebe by the first conjunct of the first sentence of (236) as a belief of Ella’s, but not the one attributed by the second. (239) below shows the effect of this accommodation on (237). First, however, we need the preliminary representation for the second sentence of (236). This representation is given in (238)

(238) $$\left\langle \left\{ \boxed{\begin{array}{c} \mathrm{q} \\ \mathrm{Ella(q)} \end{array}} \right\} ,\right.$$

$$\left. \boxed{\begin{array}{c} \mathrm{t''} \;\; \mathrm{s''} \\ \mathrm{n} \subseteq \mathrm{t''} \;\; \mathrm{t''} \subseteq \mathrm{s''} \\ \mathrm{s''{:}Att(q,} \left\{ \left\langle \mathrm{BEL}, \boxed{\left\langle \left\{ \boxed{\begin{array}{c} \mathrm{u'} \\ \mathrm{pers(u')} \\ \mathrm{male(u')} \end{array}} \right\} , \boxed{\neg \boxed{\begin{array}{c} \mathrm{t'''} \;\; \mathrm{e'''} \;\; \mathrm{w} \\ \mathrm{t'''} < \mathrm{n} \;\; \mathrm{e'''} \subseteq \mathrm{t'''} \\ \mathrm{e'''{:}steal(u',w)} \end{array}}} \right\rangle} \right\rangle \right\}) \end{array}} \right\rangle$$

Justification of (238) in the context of (237) includes, first, the justification of the presupposition introduced by the proper name *Ella*. Here we proceed as we did in the last example: assuming that the context in which this presupposition must be justified contains no more information than what is given in (237), accommodation is the only way, and it is what recipients normally do when they are confronted with a name whose referent they cannot identify by independent means.

We will assume, then, that this presupposition is accommodated and that the accommodation has yielded a discourse referent q in the main universe, which represents the referent of the name.

It is now possible to accommodate the first of the two beliefs in (237) as a belief of Ella's. The two accommodations together yield (239)

(239)
$$\boxed{\begin{array}{c}
\mathrm{t\ \ s\ \ p\ \ g\ \ q\ \ s'} \\ \hline
\mathrm{n} \subseteq \mathrm{t}\ \ \mathrm{t} \subseteq \mathrm{s}\ \ \mathrm{Phoebe(p)}\ \ \mathrm{garden\text{-}of(g,p)}\ \ \mathrm{Ella(q)}\ \ \mathrm{t} \subseteq \mathrm{s'} \\[1em]
\boxed{\begin{array}{c} \mathrm{z''} \\ \hline \mathrm{prize\text{-}z.(z'')} \\ \mathrm{poss(z'',v)} \end{array}} \ \left\langle \begin{array}{c} \forall \\ \mathrm{z''} \end{array} \right\rangle \ \boxed{\begin{array}{c} \\ \hline \mathrm{z''} = \mathrm{z} \end{array}} \\
\mathrm{s{:}Att(p,} \left\{ \begin{array}{c}
\left\langle \mathrm{BEL}, \boxed{\begin{array}{c} \mathrm{t'\ \ e'\ \ x} \\ \hline \mathrm{t'} < \mathrm{n}\ \ \mathrm{e'} \subseteq \mathrm{t'}\ \ \mathrm{man(x)} \\ \mathrm{e'{:}break\text{-}into(x,g)} \end{array}} \right\rangle \\[1em]
\left\langle \mathrm{BEL}, \boxed{\begin{array}{c} \mathrm{t'''\ \ e'''\ \ u} \\ \hline \mathrm{t'''} < \mathrm{n}\ \ \mathrm{e'''} \subseteq \mathrm{t'''}\ \ \mathrm{u} = \mathrm{x} \\ \mathrm{e'''{:}steal(u,z)} \end{array}} \right\rangle
\end{array} \right\} \mathrm{)} \\
\mathrm{s'{:}Att(q,} \left\{ \left\langle \mathrm{BEL}, \boxed{\begin{array}{c} \mathrm{t''\ \ e''\ \ x'} \\ \hline \mathrm{t''} < \mathrm{n}\ \ \mathrm{e''} \subseteq \mathrm{t'}\ \ \mathrm{man(x')} \\ \mathrm{e''{:}break\text{-}into(x',g)} \end{array}} \right\rangle \right\} \mathrm{)}
\end{array}}$$

(239) can only be regarded as an intermediate accommodation result, for we still have to deal with the presupposition generated by *he* in the last sentence of (236). This presupposition can be resolved in the secondary context given by the accommodated belief, applying the same resolution principle that we already made use of in our treatment of (220) and (228). The result of this last resolution (which takes the form of adding the discourse referent u to the universe of the DRS characterising the belief in (239) and adding "u = x′" to its conditions), and the merge with (238) which follows it, is given in (240)

$$
(240)\quad
\boxed{\begin{array}{c}
\mathrm{t\ \ s\ \ p\ \ g\ \ q\ \ s'\ \ s''}\\
\mathrm{n} \subseteq \mathrm{t}\quad \mathrm{t} \subseteq \mathrm{s}\quad \text{Phoebe(p)}\quad \text{garden-of(G,p)}\quad \text{Ella(q)}\quad \mathrm{t} \subseteq \mathrm{s'}\quad \mathrm{t} \subseteq \\[1em]
\boxed{\begin{array}{c} \mathrm{z''} \\ \text{prize-z.}(\mathrm{z''}) \\ \text{poss}(\mathrm{z''},\mathrm{v}) \end{array}}
\left\langle \begin{array}{c} \forall \\ \mathrm{z''} \end{array} \right\rangle
\boxed{\begin{array}{c} \\ \mathrm{z''} = \mathrm{z} \end{array}}
\\
\text{s:Att(p,}\left\{\begin{array}{c}
\left\langle \text{BEL,}\ \boxed{\begin{array}{c} \mathrm{t'\ \ e'\ \ x} \\ \mathrm{t'} < \mathrm{n}\quad \mathrm{e'} \subseteq \mathrm{t'}\quad \text{man(x)} \\ \text{e}'\text{:break-into(x,g)} \end{array}} \right\rangle \\[1em]
\left\langle \text{BEL,}\ \boxed{\begin{array}{c} \mathrm{t'''\ \ e'''\ \ u} \\ \mathrm{t'''} < \mathrm{n}\quad \mathrm{e'''} \subseteq \mathrm{t'''}\quad \mathrm{u} = \mathrm{x} \\ \text{e}'''\text{:steal(u,z)} \end{array}} \right\rangle
\end{array}\right\}\text{)}
\\[1em]
\text{s}'\text{:Att(q,}\left\{\begin{array}{c}
\left\langle \text{BEL,}\ \boxed{\begin{array}{c} \mathrm{t''\ \ e''\ \ x'} \\ \mathrm{t''} < \mathrm{n}\quad \mathrm{e''} \subseteq \mathrm{t''}\quad \text{man(x}'\text{)} \\ \text{e}''\text{:break-into(x}'\text{,g)} \end{array}} \right\rangle \\[1em]
\left\langle \text{BEL,}\ \boxed{\neg\ \boxed{\begin{array}{c} \mathrm{t''''\ \ e''''\ \ u'\ \ w} \\ \mathrm{t''''} < \mathrm{n}\quad \mathrm{e''''} \subseteq \mathrm{t''''}\quad \mathrm{u'} = \mathrm{x'} \\ \text{e}'''\text{:steal(u}'\text{,w)} \end{array}}} \right\rangle
\end{array}\right\}\text{)}
\end{array}}
$$

The interpretation problem on which we have focused in our discussion of (236) is closely related to one that has received a good deal of attention in the literature, especially from philosophers of language. This secon problem is known as the “Hob-Nob problem”, after the example sentence which was used by Geach to introduce the problem:

(241) Hob believes that a witch has killed Cob’s cow and Nob thinks that she has blighted Bob’s sow.

Geach pointed out that this sentence could be used truthfully in a report composed by a journalist describing the goings-on in some remote rural backwater, even if the journalist herself is persuaded that witches do not exist. This is a problem for the application of standard logical notation to the representation of truth-conditional content. For in order that the pronoun *she* in the belief attribution to Nob be bound by the “existential quantifier” *a witch* in the belief attribution to Hob, this quantifier would have to take scope over the two belief attributions. But this would, on the standard interpretation of quantification theory, imply that there are witches in the world in which Hob, Nob and the journalist live. That is something to which the journalist would under no conditions want to commit herself.And it is something to which (241) does not commit her.

30 Note that there is no need to adapt our notions of PISBAS and ISBAS to the more comprehensive repertoire of DRSs we are considering now, in which internal and external anchors have their place too. For the cases of singularity (of propositions and, by extension, of information states and CCPs to which anchoring gives rise are included in the original Def. 2. However, it is only at this point that the possibility of ISBASs containing singular information states, etc as constituents becomes essential. ADSs can now determine such singular semantic objects, and when they do, they will be subsumed by the relevant values of ASM only if those values have corresponding constituents which are singular as well. 31 Pronouns occurring in the complement of an attitude predicate or in the complement of a predicate dicendi, and which are given a de se interpretation are also sometimes called "quasi-indicators". The term was coined by Castaneda, who was the first to investigate the de se interpretation of third person pronouns closely) [ref.s]. 32 Think of Bill seeing no more than the lower part of himself, with the burning trousers, and that he thinks on the strength of what he sees: "Soon this guy's shirt will be on fire too." He hasn't seen the shirt the man is wearing but assumes, on the basis of general knowledge, that the man, if he wears the trousers he can see, will also be earing a (unique) shirt. Here poor bill may be right that hte person he is seeing doe wear a shirt, and he may also be right in thinking that that shirt will be presently on fire, but the part of his attitudes that corresponds to his shirt dos not have an anchored representation for the shirt, and a de re representation for himself as the owner of the shirt.

BIBLIOGRAPHY

[Abusch, 2004] D. Abusch. On the temporal composition of infinitives. In J.Gueron and J. Lecarme, editors, *The Syntax of Time*. MIT Press, 2004.

[Ahn and Kolb, 1990] R. Ahn and H.-P. Kolb. Discourse representation meets constructive mathematics. In L. Kalman and L. Polos, editors, *Papers from the Second Symposium on Logic and Language*, pages 105–124. Akademiai Kiadoo, Budapest, 1990.

[Asher and H.Kamp, 1989] N. Asher and H.Kamp. Self-reference, attitudes and paradox. In Gennaro Chierchia, Barbara Partee, and Raymond Turner, editors, *Properties, Types and Meaning*, volume 1. Kluwer, Dordrecht, 1989.

[Asher and Lascarides, 1998] N. Asher and A. Lascarides. Bridging. *Journal of Semantics*, pages 83–113, 1998.

[Asher and Lascarides, 2003] Nicholas Asher and Alex Lascarides. *Logics of Conversation*. Cambridge University Press, 2003.

[Asher *et al.*,] Nicholas Asher, Daniel Hardt, and Joan Busquets. Discourse parallelism, scope, and ellipsis.

[Asher, 1986] N. Asher. Belief in discourse representation theory. *Journal of Philosophical Logic*, 15:127–189, 1986.

[Asher, 1993] N. Asher. *Reference to Abstract Objects in Discourse*. Kluwer, Dordrecht, Dordrecht, 1993.

[Bach, 1981] E. Bach. On time, tense and aspect: An essay in english metaphysics. In P. Cole, editor, *Radical Pragmatics*, pages 62–81. Academic Press, New York, New York, 1981.

[Barwise and Cooper, 1993] J. Barwise and R. Cooper. Extended Kamp notation. In P. Aczel, D. Israel, Y. Katagiri, and S. Peters, editors, *Situation Theory and its Applications, Volume III*, chapter 2, pages 29–54. CSLI, 1993.

[Barwise *et al.*, 1991] J. Barwise, J. M. Gawron, G. Plotkin, and S. Tutiya. *Situation Theory and its Applications, volume II.* CSLI and University of Chicago, 1991.

[Beaver, 1997] David Ian Beaver. Presupposition. In Johan van Benthem and Alice ter Meulen, editors, *Handbook of Logic and Language*, pages 939–1008. Elsevier, Amsterdam, 1997.

[Beaver, 2004] David Ian Beaver. Accomodating topics. In H. Kamp and B. H. Partee, editors, *Context-Dependence in the Analysis of Linguistic Meaning*, pages 79–90. Elsevier, Amsterdam, 2004.

[Benthem, 1983] J.F.A.K. van Benthem. *The Logic of Time.* Reidel, Dordrecht, Dordrecht, 1983.

[Berman and Hestvik, 1994] S. Berman and A. Hestvik. Principle b, drt and plural pronouns. to appear in: Arbeitspapiere des sonderforschungsbereichs 340, IMS Stuttgart, Germany, 1994.

[Bierwisch, 1983] M. Bierwisch. Semantische und konzeptuelle repräsentationen lexikalischer einheiten. In R. Ruzicka and W. Motsch, editors, *Untersuchungen zur Semantik, Studia Grammatika 22*, pages 61–99. Akademie Verlag, Berlin, 1983.

[Bonevac, 1987] Daniel Bonevac. *Deduction.* Mayfield, 1987.

[Bos *et al.*, 1994] J. Bos, E. Mastenbroek, S. McGlashan, S. Millies, and M. Pinkal. A compositional DRS-based formalism for NLP-applications. In *International Workshop on Computational Semantics*, Tilburg, 1994.

[Bos *et al.*, 1995] J. Bos, A.-M. Mineur, and P. Buitelaar. Bridging as coercive accommodation. In *CLNLP*, Edinburgh, 1995.

[Briscoe *et al.*, 1993] E. Briscoe, V. de Paiva, and A. Copestake, editors. *Inheritance, Defaults and the Lexicon.* Cambridge University Press, 1993.

[Caenepeel, 1989] Mimo Caenepeel. *Aspect, Temporal Ordering and Perspective in Narrative Fiction.* PhD thesis, University of Edinburgh, 1989.

[Chierchia, 1991] G. Chierchia. Anaphora and dynamic binding. *Linguistics and Philosophy*, 15(2):111–183, 1991.

[Chierchia, 1993] G. Chierchia. Questions with quantifiers. *Natural Language Semantics*, 1:181–234, 1993.

[Chomsky, 1981] N. Chomsky. *Lectures on Government and Binding.* Foris, Dordrecht, 1981.

[Clark, 1997] H. Clark. Bridging. In P. N. Johnson-Laird and P. C. Wason, editors, *Thinking: Readings in Cognitive Science*, pages 411–420. Cambridge University Press, 1997.

[Cooper *et al.*, 1990] R. Cooper, K. Mukai, and J. Perry. *Situation Theory and its Applications, volume I.* CSLI and University of Chicago, 1990.

[Cooper, 1975] R. Cooper. *Montague's Semantic Theory and Transformational Syntax.* PhD thesis, University of Massachusetts, Amherst, 1975.

[de Swart, 1999] H. de Swart. Negation and the temporal structure of narrative discourse. *Journal of Semantics*, pages 1–42, 1999.

[Dekker, 1993] P. Dekker. *Transsentential Meditations.* PhD thesis, Department of Philosophy, University of Amsterdam, 1993.

[Gabbay and Finger, 2000] Mark Reynolds Dov Gabbay and Marcelo Finger. *Temporal Logic*, Volume II. Oxford, 2000.

[Gabbay and Reynolds, 1994] Ian Hodkinson Dov Gabbay and Mark Reynolds. *Temporal Logic*, Volume I. Oxford, 1994.

[Dowty, 1979] D.R Dowty. *Word Meaning and Montague Grammar.* Reidel, Dordrecht, Dordrecht, 1979.

[Eberle, 1997] Kurt Eberle. Flat underspecified representation and its meaning for a fragment of german. Technical report, Sonderforschungsbereich 340, Report, 1997.

[Eijck and Kamp, 1997] J. van Eijck and H. Kamp. Representing discourse in context. In J. van Benthem and A. ter Meulen, editors, *Handbook of Logic an Language*, pages 179–237. Elsevier, Amsterdam, 1997.

[Engdahl, 1980] E. Engdahl. *The Syntax and Semantics of Questions in Swedish.* PhD thesis, University of Massachusetts at Amherst, 1980. Distributed by the GLSA, Department of Linguistics, University of Massachusetts at Amherst.

[Fabricius-Hansen, 1980] Catherine Fabricius-Hansen. Lexikalische Dekomposition, Bedeutungspostulate und w i e d e r. Ein Beitrag zu einer Montague- Grammatik des Deutschen. In Dieter Katovsky, editor, *Perspektiven der lexikalischen Semantik*, pages 26 – 40. Kohlhammer, 1980.

[Fabricius-Hansen, 1983] Cathrine Fabricius-Hansen. Wieder ein w i e d e r? Zur Semantik von w i e d e r. In Rainer Bäuerle, Christoph Scharze, and Arnim von Stechow, editors, *Meaning, Use and Interpretation of Language*, pages 97–120. Berlin, New York: de Gruyter, 1983.

[Fernando, 1992] Tim Fernando. Transition systems and dynamic semantics. In *Logics in AI, LNCS 633.* Springer-Berlin, Berlin, 1992.

[Fernando, 1994] Tim Fernando. What is a drs? In *First International Workshop on Computational Semantics*, Tilburg, The Netherlands, 1994.

[Fernando, 2001a] Tim Fernando. Conservative generalized quantifiers and presupposition. In *Proceedings, Eleventh Semantic and Linguistic Theory conference (SALT XI)*, pages 172–191, New York University, 2001.

[Fernando, 2001b] Tim Fernando. A type reduction from proof-conditional to dynamic semantics. *Journal of Philosophical Logic*, 30(2):121–153, 2001.

[Fodor, 1975] J. Fodor. *The Language of Thought.* Harvard University Press, 1975.

[Frank and Reyle, 1995] A. Frank and U. Reyle. Principle based semantics for hpsg. In *Proceedings of the 7th Conference of the EACL, March 27-31, Dublin*, pages 9–16, 1995.

[Gallin, 1975] D. Gallin. *Intensional and Higher Order Logic; with Applications to Montague Semantics.* North Holland, Amsterdam, 1975.

[Gamut, 1991] L.T.F. Gamut. *Language, Logic and Meaning, Part 2.* Chicago University Press, Chicago, 1991.

[Gärdenfors, 1988] P. Gärdenfors. *Knowledge in Flux: Modelling the Dynamics of Epistemic States.* MIT Press, Cambridge Mass, 1988.

[Geach, 1962 (Third revised edition: 1980)] P.T. Geach. *Reference and Generality : An Examination of Some Medieval and Modern Theories.* Cornell University Press, Ithaca, 1962 (Third revised edition: 1980).

[Genabith and Crouch, 1999] J. van Genabith and R. Crouch. Dynamic and underspecified semantics for lfg. In Mary Dalrymple, editor, *Semantics and Syntax in Lexical Functional Grammar: The Resource Logic Approach*, pages 209–260. MIT Press, Cambridge, Massachusetts, 1999.

[Geurts and van der Sandt, 1999] Bart Geurts and Rob van der Sandt. Presuppositions and backgrounds. In *Proceedings of the 11th Amsterdam Colloquium.* University of Amsterdam, 1999.

[Geurts, 1997] B. Geurts. Good news about the description theory of names. *Journal of Semantics*, 14:319–348, 1997.

[Geurts, 1999] Bart Geurts. *Presuppositions and Pronouns.* Elsevier Science, 1999.

[Goldblatt, 1992 (first edition 1987)] R. Goldblatt. *Logics of Time and Computation, Second Edition, Revised and Expanded*, volume 7 of *CSLI Lecture Notes.* CSLI, Stanford, 1992 (first edition 1987). Distributed by University of Chicago Press.

[Grice, 1961] P. Grice. The causal theory of perception. *Aristotelian Society Proceedings, Supplementary Volume*, 35, 1961. Reprinted in: P. Grice: Studies in the Way of Words.

[Grimm and Merrill, 1988] R. Grimm and D. Merrill, editors. *Contents of Thought.* University of Arizona Press, 1988.

[Groenendijk and Stokhof, 1990] J. Groenendijk and M. Stokhof. Dynamic montague grammar. In L. Kalman and L. Polos, editors, *Papers from the Second Symposium on Logic and Language*, pages 3–48. Akademiai Kiadoo, Budapest, 1990.

[Groenendijk and Stokhof, 1991] J. Groenendijk and M. Stokhof. Dynamic predicate logic. *Linguistics and Philosophy*, 14:39–100, 1991.

[Hans and Reyle,] Kamp Hans and Uwe Reyle. From discourse to logic. Volume II.

[Hardt, 1992] D. Hardt. An algorithm for vp ellipsis. In *Proceedings 30th ACL, Newark*, pages 9–14, 1992.

[Harel, 1984] D. Harel. Dynamic logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, pages 497–604. Reidel, Dordrecht, 1984. Volume II.

[Heim, 1982] I. Heim. *The Semantics of Definite and Indefinite Noun Phrases*. PhD thesis, University of Massachusetts, Amherst, 1982.

[Heim, 1983] I. Heim. On the projection problem for presuppositions. *Proceedings of the West Coast Conference on Formal Linguistics*, 2:114–126, 1983.

[Hinrichs, 1986] E. Hinrichs. Temporal anaphora in discourses of english. *Linguistics and Philosophy*, 9(1):63–81, 1986.

[Hobbs, 1990] Jerry R Hobbs. The coherence and structure of discourse. In *Literature and cognition*. CSLI, 1990. CSLI Lecture notes, 21.

[Hodges, 2001] Wilfrid Hodges. Elementary predicate logic. In Dov Gabbay, editor, *Handbook of Philosophical Logic*, volume I. Reidel, second edition, 2001.

[Kadmon, 1987] N. Kadmon. *On Unique and Non-Unique Reference and Asymmetric Quantification*. PhD thesis, University of Massachusetts at Amherst, 1987.

[Kalish and Montague, 1964] Kalish and Montague. *Logic. Techniques of Formal Reasoning*. Harcourt, Bran & World, New York, 1964.

[Kamp and Reyle, 1991] H. Kamp and U. Reyle. A calculus for first order discourse representation structures. Arbeitspapiere des Sonderforschungsbereichs 340 16, IMS Stuttgart, Germany, 1991.

[Kamp and Reyle, 1993] H. Kamp and U. Reyle. *From Discourse to Logic*. Kluwer, Dordrecht, 1993.

[Kamp and Rohrer, 1983a] H. Kamp and C. Rohrer. Tense in texts. In Bäuerle, Schwarze, and Von Stechow, editors, *Meaning, Use and Interpretation of Language*, pages 250–269. De Gruyter, Berlin, 1983.

[Kamp and Rohrer, 1983b] Hans Kamp and Christian Rohrer. Temporal Reference in French. ms. Stuttgart, 1983.

[Kamp and Roßdeutscher, 1994] H. Kamp and A. Roßdeutscher. Remarks on lexical structure and drs-construction. (ed) H. Schnelle, Lexical Semantics, special issue of Theoretical Linguistics, 1994.

[Kamp and Schiehlen, 2002] Hans Kamp and Michael Schiehlen. *How to say WHEN it happens*, chapter Temporal Location in Natural Language. Niemeyer, 2002.

[Kamp, 1971] Hans Kamp. Formal properties of "NOW". *Theoria*, 37:227 – 273, 1971.

[Kamp, 1979] H. Kamp. Events, instants and temporal reference. In Bäuerle, Egli, and Von Stechow, editors, *Semantics from Different Points of View*. Springer, Berlin, 1979.

[Kamp, 1981a] H. Kamp. A theory of truth and semantic representation. In J. Groenendijk et al., editors, *Formal Methods in the Study of Language*. Mathematisch Centrum, Amsterdam, 1981.

[Kamp, 1981b] Hans Kamp. Evénements, representation discursive et reference temporelle. *Langages*, 64:39–64, 1981.

[Kamp, 1988] H. Kamp. Comments on stalnaker, belief attribution and context. In R. Grimm and D. Merrill, editors, *Contents of Thought*. University of Arizona Press, 1988.

[Kamp, 1990] H. Kamp. Prolegomena to a structural account of belief and other attitudes. In C. A. Anderson and J. Owens, editors, *Propositional Attitudes—The Role of Content in Logic, Language, and Mind*, chapter 2, pages 27–90. University of Chicago Press and CSLI, Stanford, 1990.

[Kamp, 1999] Hans Kamp. The representation of attitudes and their causal connections. ms. university of Stuttgart, 1999.

[Kamp, 2001a] Hans Kamp. The importance of presupposition. In Christian Rohrer, Antje Rossdeutscher, and Hans Kamp, editors, *Liguistic Form and its Computation*. CSLI-Publications, Standord, 2001.

[Kamp, 2001b] Hans Kamp. Presupposition computation and presupposition justification. In Myriam Bras and Laure Vie, editors, *Pragmatic and Semantic Issues in Discourse and Dialogue*. Elsevier, 2001.

[Kamp, 2003] H. Kamp. Einstellungszustände und -zuschreibungen in der drt. In Ulrike Haas-Spohn, editor, *Intentionalität zwischen Subjektivit"at und Weltbezug*, pages 209–289. mentis, Paderborn, 2003.

[Kaplan and Bresnan, 1982] R.M. Kaplan and J. Bresnan. Lexical functional grammar. In J. Bresnan, editor, *The mental representation of grammatical relations*, pages 173–281. MIT Press, Cambridge Mass, 1982.

[Kaplan, 1969] D. Kaplan. Quantifying in. In D. Davidson and J. Hintikka, editors, *Words and Objections: Essays on the Work of W. V. Quine*. Dordrecht: Reidel, 1969.

[Kaplan, 1989] D. Kaplan. Demonstratives. In H. Wettstein J. Almog, J Perry, editor, *Themes form Kaplan*. Oxford University Press, 1989.

[Karttunen, 1976] L. Karttunen. Discourse referents. In J. McCawley, editor, *Syntax and Semantics 7*, pages 363–385. Academic Press, 1976.

[Keenan and Westerstahl, 1997] E. Keenan and D. Westerstahl. Generalized quantifiers in liguistics and logic. In J. van Benthem and A. ter Meulen, editors, *Handbook of Language and Logic*, pages 837–893. Elsevier, 1997.

[Keenan, 1992] E. Keenan. Beyond the frege boundary. *Linguistics and Philosophy*, pages 199–221, 1992.

[Keisler, 1970] H. Jerome Keisler. Logic with the quantifier 'there exist uncoutably many'. *Annals of Mathematical Logic*, 1:1–93, 1970.

[Kohlhase *et al.*, 1996] M. Kohlhase, S. Kuschert, and M. Pinkal. A type-theoretic semantics for λ-drt. In P. Dekker and M. Stokhof, editors, *Proceedings of the Tenth Amsterdam Colloquium*. ILLC, University of Amsterdam, 1996.

[Koons, 1988] Koons. Deduction system for drt. (ms.) Austin, Texas, 1988.

[Krause, 2001] Peter Krause. *Topics in Presupposition Theory*. PhD thesis, Universität Stuttgart, 2001.

[Kripke,] S.A. Kripke. Presuppositions and anaphora: Remarks on the formulation of the projection problem. Princeton University.

[Kripke, 1972] S.A. Kripke. Naming and necessity. In D. Davidson and G. Harman, editors, *Semantics of Natural Language*, pages 253–355. Reidel, Dordrecht, 1972.

[Kripke, 1979] S. Kripke. A puzzle about belief. In A. Margalit, editor, *Meaning and Use*. Reidel, Dordrecht, 1979.

[Lascarides and Asher, 1993] Alex Lascarides and Nicholas Asher. Temporal interpretation, discourse relations, and common sense entailment. *Linguistics and Philosophy*, 16:437–49, 1993.

[Lascarides and Copestake, 1999] Alex Lascarides and Ann Copestake. Default representation in constraint-based frameworks. *Computational Linguistics*, 1999.

[Lasersohn, 1995] Peter Lasersohn. *Plurality, Conjunction and Events*. Kluwer Academic Publishers, 1995.

[Lasnik, 2003] H. Lasnik. *Minimalist Investigations in Linguistic Theory*. Routledge Ltd, 2003.

[Lewis, 1975] D. K. Lewis. Adverbs of quantification. In E. Keenan, editor, *Formal Semantics of Natural Language*. Cambridge University Press, Cambridge, 1975.

[Link, 1983] G. Link. The logical analysis of plurals and mass terms. In R. Bäuerle, C. Schwarze, and A. von Stechow, editors, *Meaning, Use and Interpretation of Language*, pages 303–323. Walter de Gruyter, Berlin, 1983.

[Loar, 1987] B. Loar. Subjective intentionality. *Philosophical topics*, 15:89–124, 1987.

[Loar, 1988] B. Loar. Social content and psychological content. In R. Grimm and D. Merill, editors, *Contents of Thought*, pages 99–110. University of Arizona Press, 1988.

[Martin-Löf, 1984] P. Martin-Löf. *Intuitionistic Type Theory*. Bibliopolis, 1984.

[Moens and Steedman, 1988] M. Moens and M. Steedman. Temporal ontology and temporal reference. *Computational Linguistics*, 14(2):15–28, 1988.

[Montague, 1973] R. Montague. The proper treatment of quantification in ordinary english. In J. Hintikka e.a., editor, *Approaches to Natural Language*, pages 221–242. Reidel, 1973.

[Muskens *et al.*, 1997] R. Muskens, J. van Benthem, and A. Visser. Dynamics. In J. van Benthem and A. ter Meulen, editors, *Handbook of Logic an Language*, pages 587–648. Elsevier, Amsterdam, 1997.

[Muskens, 1996] R. Muskens. Combining montague semantics and discourse representation. *Linguistics and Philosophy*, 19:143–186, 1996.

[Ogihara, 1999] T. Ogihara. Double-access sentences generalized. In *Proceedings from Semantics and Linguistic Theory IX (SALT9)*. CLC Publications, 1999.

[Partee, 1984] B.H. Partee. Nominal and temporal anaphora. *Linguistics and Philosophy*, 7, 1984.

[Pollard and Sag, 1994] C. Pollard and I. Sag. *Head-Driven Phrase Structure Grammar*. CSLI Lecture Notes. CSLI, Stanford, 1994. Distributed by University of Chicago Press.

[Pratt, 1976] V. Pratt. Semantical considerations on Floyd–Hoare logic. *Proceedings 17th IEEE Symposium on Foundations of Computer Science*, pages 109–121, 1976.

[Quine, 1956] W.V.O Quine. Quantifiers and propositional attitudes. *Journal of Philosophy*, 1956.

[Quine, 1961] W.V.O Quine. Reference and modality. In *From a Logical Point of View*. New York: Harper and Row, 1961. Reprinted in Linsky: Reference and Modality, London: Oxford University Press, 1971.

[Ranta, 1995] Aarne Ranta. *Type-theoretical Grammar*. Oxford University Press, 1995.

[Reichenbach, 1947] H. Reichenbach. *Elements of Symbolic Logic*. Macmillan, London, 1947.

[Reinhard, 1989] Reinhard. Deduktionen auf diskursrepräsentationsstrukturen. Studienarbeit, IMS Stuttgart, Germany, 1989.

[Reyle and Frey, 1983] U. Reyle and W. Frey. A prolog implementation of lexical functional grammar. In *IJCAI 83, Karlsruhe, Germany*, pages 693–695, 1983.

[Reyle and Gabbay, 1994] U. Reyle and D. Gabbay. Direct deductive computation on discourse represenation structures. *Linguistics and Philosophy*, 17:343–390, 1994.

[Reyle and Rossdeutscher, 2001] Uwe Reyle and Antje Rossdeutscher. Temporal underspecification in discourse. In Christian Rohrer, Antje Rossdeutscher, and Hans Kamp, editors, *Linguistic Form and its Computation*. CSLI, 2001.

[Reyle *et al.*, 2000] Uwe Reyle, Antje Rossdeutscher, and Hans Kamp. Ups and downs in the theory of temporal reference. In Uwe Reyle, editor, *Presuppositions and Underspecification in the Computation of Temporal and other Relations in Discourse*. Arbeitsberichte des Sonderfoschungsbereichs 340, Stuttgart/Tübingen, Nr.164, 2000.

[Reyle *et al.*, 2003] Uwe Reyle, Antje Rossdeutscher, and Hans Kamp. Ups and downs in the theory of temporal reference. ms. University of Stuttgart, 2003.

[Reyle, 1993] U. Reyle. Dealing with ambiguities by underspecification: Construction, representation and deduction. *Journal of Semantics*, 10:123–179, 1993.

[Rooth, 1992] Mats Rooth. A theory of focus interpretation. *Natural Language Semantics*, 1:75–116, 1992.

[Rooth, 2005] Mats Rooth. Topic accents on quantifiers. In Greg Carlson and Jeffrey Pelletier, editors, *Reference and Quantification: the Partee Effect*. CSLI Publications, 2005.

[Rooy, 1997] Robert Van Rooy. *Attitudesw and Changing Contexts*. PhD thesis, IMS, University of Stuttgart, 1997.

[Roßdeutscher and Reyle, 2000] Antje Roßdeutscher and Uwe Reyle. Constraint-based bottom up discourse interpretation. In Uwe Reyle, editor, *Presuppositions and Underspecification in the Computation of Temporal and other Relations in Discourse*. Arbeitsberichte des Sonderfoschungsbereichs 340, Stuttgart/Tübingen, Nr.164, 2000.

[Roßdeutscher, 2000] Antje Roßdeutscher. *Lexikalisch gesttzte formale Textinterpretation*. Arbeitsberichte des Sonderfoschungsbereichs 340, Stuttgart/Tübingen, Nr.157, 2000.

[Russell, 1905] B. Russell. On denoting. *Mind*, 14:479–493, 1905.

[Sandström, 1993] Gorel Sandström. *When-Clauses and the Temporal Interpretation of Narrative Discourse*. PhD thesis, University of Umea, 1993.

[Saurer, 1993] W. Saurer. A natural deduction system of discourse representation theory. *Journal of Philosophical Logic*, 22(3):249–302, 1993.

[Schiehlen, 1999] Michael Schiehlen. *Semantikkonstruktion*. PhD thesis, Universität Stuttgart, 1999.

[Schiehlen, July 2002] Michael Schiehlen. Ellipsis resolution with scope underspecification. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics,*, Philadelphia PA, July 2002.

[Sedogbo, 1988] C. Sedogbo. Sylog: A drt system in prolog. In V. Dahl and P. Saint-Dizier, editors, *Natural Language Understanding and Logical Programming II*, pages 185–201. Elsevier Science Publishers, 1988.

[Seizmair, 1996] Michael Seizmair. Nicht-fregeische quantifikation und dynamik. Diplomarbeit, University of Stuttgart, 1996.

[Seuren, 1986] P. Seuren. *Discourse Semantics*. Blackwell, Oxford, 1986.

[Soames, 1984] S. Soames. Presupposition. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, pages 553–616. Reidel, 1984. Volume IV.

[Stalnaker, 1972] R. Stalnaker. Pragmatics. In D. Davidson and G. Harman, editors, *Semantics of Natural Language*, pages 380–397. Reidel, 1972.

[Stalnaker, 1974] R. Stalnaker. Pragmatic presuppositions. In M.K. Munitz and P.K. Unger, editors, *Semantics and Philosophy*. New York University Press,, 1974.

[Stalnaker, 1979] R. Stalnaker. Assertion. In P. Cole, editor, *Syntax and Semantics Vol.9: Pragmatics*, pages 315–332. Academic Press, New York, 1979.

[Stalnaker, 1988] R. Stalnaker. Belief attribution and context. In R. Grimm and D. Merrill, editors, *Contents of Thought*. University of Arizona Press, 1988.

[Stalnaker, 1990] R. Stalnaker. Narrow content. In C. Antony Anderson and joseph Owens, editors, *Propositional Attitudes: the Role of Content in Logic, Language and Mind*. CSLI, 1990. Reprinted in: R. C. Stalnaker (1999): Context and Content.

[Stechow, 1996] Arnim von Stechow. The different readings of "wieder" (again): A structural account. *Journal of Semantics*, 1996.

[Steedman, 2001] M. Steedman. *The Syntactic Process*. Bradford Books, MIT Press, 2001.

[Stone, 1998] Matthiew Stone. *Modality in Dialogue: Planning, Pragmatics and Computation*. PhD thesis, University of Pennsylvania, 1998.

[Strawson, 1950] P.F. Strawson. On referring. *Mind*, 59:320–344, 1950.

[Strawson, 1964] Strawson. Intention and convention in speech acts. *Philosophical Review*, 1964.

[Sundholm, 1986] G. Sundholm. Proof theory and meaning. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic III*, pages 471–506. Kluwer, Dordrecht, 1986.

[Sundholm, 2001] G. Sundholm. Systems of deduction. In D. Gabbay, editor, *Handbook of Philosophical Logic II. Second Edition*. Kluwer, Dordrecht, 2001.

[Thomason, 2002] R. Thomason. Combinations of tense and modality. In D. Gabbay, editor, *Handbook of Philosophical Logic VII*. Kluwer, Dordrecht, 2002.

[Turner, 1988] R. Turner. A theory of properties. *The Journal of Symbolic Logic*, 54, 1988.

[van der Does and van Eijck (eds.), 1991] J. van der Does and J. van Eijck (eds.). *Generalized Quantifier Theory and Applications*. Dutch Network for Language, Logic and Information, 1991.

[van der Sandt, 1992] R.A. van der Sandt. Presupposition projection as anaphora resolution. *Journal of Semantics*, 9:333–377, 1992. Special Issue: Presupposition, Part 2.

[Vermeulen, 1995] C.F.M. Vermeulen. Merging without mystery, or: Variables in dynamic semantics. *Journal of Philosophical Logic*, 35:405–450, 1995.

[von Fintel, 1994] Kai von Fintel. *Restrictions on Quantifier Domains*. PhD thesis, GLSA, UMass Amherst, 1994.

[Webber, 1988] B.L. Webber. Tense as discourse anaphor. *Computational Linguistics*, 14(2):61–73, 1988.

[Westerstahl, 1989a] D. Westerstahl. Quantifiers in formal and natural languages. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic, volume IV*, pages 1–131. Reidel, 1989.

[Westerstahl, 1989b] D. Westerstahl. Topics in generalized quantifiers. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic, Volume IV*, pages 1–131. Reidel, 1989.

[Winter, 2002] Yoad Winter. *Flexibility Principles in Boolean Semantics. The Interpretation of Coordination, Plurality, and Scope in Natural Language*. The MIT Press, 2002.

[Zeevat *et al.*, 1987] H. Zeevat, E. Klein, and J. Calder. Unification categorial grammar. In N. Haddock, E. Klein, and G. Morrill, editors, *Categorial Grammar, Unification Grammar and Parsing*, pages 195–222. Centre for Cognitive Science, University of Edinburgh, 1987.

[Zeevat, 1989] H. Zeevat. A compositional approach to discourse representation theory. *Linguistics and Philosophy*, 12:95–131, 1989.

[Zeevat, 1992] H. Zeevat. Presupposition and accommodation in update semantics. *Journal of Semantics*, 9(4):379–412, 1992. Special Issue: Presupposition, Part 2.

INDEX

$\Box$, 156
$\Diamond$, 156
$\Im$, 147
Λ, 149
α-congruence, 10
$\mathcal{W}$ algorithm, 57
  completeness, 61
  soundness, 61
λ-abstraction, 162
λ-algebra, 34, 37
λ-definability, 44
λ-model, 34, 37
λ-terms, 3
$\models_r$, 306
$\rightarrow_h$, 30
$\subset_{\{y\}}$, 148
$\uplus$, 140
Y – fixed point combinator, 39

abduction, 311
abstraction, 184
abutment, 211
Acc_K, 302
accessibility, 135, 274, 286, 292, 302
Accessibility
  acc, 147
accessibility
  local, 292
accessible, 155
accommodation
  global, 314, 315
  intermediate, 312, 313
ADEX(K″,K′,K), 297
adjunction
  emmediate, 297
  expansion, 297
Aktionsart, 372
alphabetic variant, 252
anaphora, 127, 143
anchor
  internal, 333
antecedent
  formation, 185
applicative structure, 34
architecture, 131, 288, 325
architecture
  bottom-up, 139
  top-down, 138
argument, 249
aspect, 200
assignment, 148, 152
assignment
  updated, 153
ATCON_L, 297
atom, 187
attitude
  ascription, 339
  attribution, 339
  change, 371
  description, 339
  description set, 340, 343
  shared, 382
attitude description set
  well-founded, 348
attitudinal hierarchy, 351
axiomatisable, 179
axiomatizability, 194

Barendregt, H. P., 2
binding
  local, 304
  primary, 171
  quantificational, 171
  secondary, 171
  unselective, 165

D. Gabbay and F. Guenthner (eds.),
Handbook of Philosophical Logic, Volume 15, 395–401.
DOI 10.1007/978-94-007-0485-5,

binding constraint, 317, 318, 320
BNF syntax, 2
boolean values, 41
bridging, 271
BV(K), 146

calendar concept, 235
calendar predicate, 235
calendar predicates, 209
CCP, 145
CCP
    base, 161
    distributive, 161
    regular, 161
change of bound variables, 10
Church, A., 9
Church-Rosser Theorem, 22
closed under equality, 49
coherence, 261
cohesion, 127
combinatory algebra, 35
common ground, 371
commuting rules, 27
complete, 187
completeness, 16
condition
    atomic, 146
    complex, 146
    intensional, 156
conditional, 272
conditional proof, 255
congruence, 19
conjunction, 285
conjunction
    dynamic, 285
conservative, 172
consistency, 15
contains, 251
content
    descriptive, 269
    narrow, 356, 357
    subsumption, 352
    wide, 356, 357
context, 125, 128, 141, 143, 259
context
    global, 260
    local, 260, 292, 301
    local
        potential, 299
    primary, 360
    secondary, 360
context change potential, 145, 155, 159
context change potential
    extensional, 162
context update, 128
contexts, 5
contractum, 21
convention
    notational, 367
convertibility, 6
coordination, 285
CP, 254
curried function, 6

de Bruijn, N., 12
de dicto, 357
de re, 364, 365, 369
de se, 365, 368, 369
descendants, 28
detachment, 252
DI, 254
diamond property, 22
direct proof, 252
direct proof
    rule of, 251
discourse referent, 133, 145
discourse referent
    anaphoric, 299
    bound, 146
    free, 146
    indexical, 203, 331, 353
    neutral, 189
    plural, 189
    set denoting, 181
    singular, 189
discourse relations, 207

discourse representation theory, DRT, 130
disjunction introduction, 254
DNE, 253
doneky sentence
  strong reading, 168
  weak reading, 168
donkey sentence, 135
donkey sentences, 129
double negation elimination, 253
DRS, 131, 133
DRS
  condition, 145
  improper, 154
  language, 145
  proper, 146, 299
  pure, 147
  quasi-, 297
  sequencing, 153
  vocabulary, 145
DRS condition, 133
DRS condition
  atomic, 134, 297
  complex, 134
  conditional, 134
  conjunctive, 285
  implicational, 164
  negative, 134
  quantificational, 134
  reducible, 138
DRS construction, 137
duplex condition, 167
duplex conditions, 171
Dynamic Logic, 151
dynamic semantics, 125, 131
dynamics and representational account of meaning, 127

economy, 311
elaboration, 207
embedding, 148
embedding
  partial, 147
  verifying, 147, 157
embeddings
  verifying, 148
equality
  extensional, 14
  intensional, 14
essential undecidability, 49
EU(K′,K), 301
event, 200
eventuality, 199
existence-and-uniqueness, 263
explanation, 207
extended flashback, 214
extended polynomials, 53
extension, 148

File Change Semantics, 259
first-order, extensional DRS language, 145
fixed point, 7
fixed point combinator, 39
Fixed Point Theorem, 7
free, 187
FV(K), 146

Gödel number — #, 48
Gödel, K., 48
generalised quantifiers, 163
Girard, J.-Y., 54
grand reduction, 23
granularity, 230

head, 297
head normal form, 30
  principal, 31
higher order logic
  intensional, 162
Hindley, R., 53
Hindley–Rosen Lemma, 27
Hob-Nob problem, 386

IMADEX(K), 297
imperfect paradox, 204
imperfective paradox, 204
incompatibility, 15

inductive definition, 2
inference, 306, 309, 311
information
    presupposition-free, 303
information change potential, 125
information state, 157, 160
information state
    expressed by, 161
    ordering, 160
information states, 153
initial functions, 45
initial labelling, 29
intensional model, 155
intensionality, 344
intentionality, 344
interpretation function, 147
IPA(x,K), 305
ISBAS, 349
isomorphic, 187

justification
    global, 272
    local, 272, 284, 285, 304

labelled terms, 28
leftmost redex, 30
Leibniz Law, 13
location time, 200
LocCon$_K$(K″,K), 303
LocConK(K″,K′,K), 303
logic
    classical, 150
    first order, 150
    interval, 197
    point, 197
logical omniscience, 344

meaning postulate, 232, 248
merge
    consistent, 160
merge operation, 140
Milner, R., 56
minimalisation, 46
Mitschke's Theorem, 27
mode indicator, 330, 343
model
    lattice theoretic, 187
    temporal, 244
modus ponens
    generalised, 252
modus tollendo ponens, 254
Modus Tollens, 257
Montague Grammar, 136
MTP, 254

Name, 145
necessity
    historical, 243
negation, 282
NEU, 254
Newman's Lemma, 25
non-empty universe, 254
non-trivial subset, 49
normal form, 16
notion of reduction, 20
numeral system, 43
    adequacy, 44
    normal, 43
    standard, 43
numerals
    Church, 44
    standard, 42

operator
    aspectual, 205

PA(x,K), 305
pairs, 41
partial functions, 47
past
    simple, 198
Peano Arithmetic, 194
perspective time, 199, 215
PISBAS, 349
PLC(K), 299
plural, 181
plural morphology, 189
PR$_L^-$, 297

PR_L, 300
pragmatics, 203
PRCON_L^-, 297
PRCON_L, 300
predecessor, 43
predicate constants, 145
predicate logic
  many sorted, 233
preliminary representations, 315
PRES(K′,K), 296, 297
PRES(K), 298
PRESR(K), 299
PRESRED, 301
PRESRED(K), 298
presupposition, 131, 259, 297
presupposition
  again, 278, 282, 320
  accommodation, 261, 270, 273, 277, 306, 308, 311–313, 315
  accomodation, 261
  anaphoric, 299, 300
  cancellation, 307
  cancelling, 284
  computation, 325
  definite descriptions, 262
  factive, 275
  justification, 141, 267, 270, 272, 274, 304, 309, 325
  maximality, 310
  projcetion, 304
  projection, 272, 274
  pronoun, 262
  propositional, 300
  reduction, 298
  referential, 160
  satisfaction, 284
  trigger, 275
presupposition trigger, 260
presuppositon
  anaphoric, 265
primitive recursion, 45
principal type scheme, 53
Priorean logic, 196
PROG, 205
progressive, 205
projection functions, 42
Projection Problem, 260
pronoun
  possessive, 266
proper, 301
proper name, 142, 273
proportion problem, 166
proposition, 156, 158, 160
proposition
  expressed by, 161
propositions, 153

quantification
  universal, 272
Quantificational Dynamic Logic, 151
quantifier
  complex, 164
  existential, 149, 164
  non-Fregean, 177
  polyadic, 177
  universal, 164, 275
quantifying in, 241
quantifying-in, 130, 137

RAA, 254
RDP, 251
recursive functions, 46
recursive separability, 49
redex, 18
reductio ad absurdum, 254, 256
reduction, 18
  standard, 32
Ref, 145
reference point, 199
reference time, 215
referential transparency, 13
Rel^n, 145
relation
  compatible, 19
  reduction, 19
relation constants, 145

representation
    preliminary, 297
residual, 28
resolution
    anaphora, 271
result, 207
Reynolds, J., 54
rhetorical relations, 207
rigidity, 244
root, 297
rule
    α, 10
    β, 6
    δ, 26
    η, 10, 14
    ξ, 6
    ext, 14

Scott's Theorem, 49
Scott, D., 33
Second Fixed Point Theorem, 48
second order, 194
self-identity, 254
semantic interpretation
    dynamic , 159
semantics, 33
semantics
    intensional, 153, 155, 344
semilattice
    upper, 187
sequence of tense, 371
Show-line, 250
SI, 254
SILC(K′), 303
singular term, 269
Situation Semantics, 148
SoI, 254
Standardisation Theorem, 33
state, 200
state
    quantificational, 226
storage, 316, 317
Sub-DRS, 134
subordination, 146
subordination
    immediate $<$, 146
substitution, 9
Substitution Lemma, 13
substitution of identicals, 254
subterm, 4
successor, 42
Summation, 183
Sym, 145
syntax, 126

T_E, 298
tense and aspect, 196
tense logic, 196
term
    polymorphic λ-calculus, 54
    simply typed, 51
term model, 38
test for zero, 43
theory of tense
    two-dimensonal, 199
TILC(K′,K), 303
time, 197
tree
    decorated, 298
truth, 127, 135, 148, 149, 156
type
    closure, 59
    ground, 50
    intersection, 62
    most general, 57
    simple, 50
Type Theory
    constructive, 153
    intuitionistic, 153
typing
    Church, 50
    Curry, 50

unifier, 58
    most general, 58
universe
    extended, 301
    main, 149

update, 131
utterance time, 199

validity, 234, 249
van Bakel, S., 63
variable convention, 11
variables
    bound, 4
    free, 4
verb
    change-of-state, 373
verification, 148
verification
    r-verification, 306
    DRS, 149
    DRS condition, 149

weak diamond property, 25
weak extensionality, 6
weak head normal form, 14, 32
world
    possible, 155, 156
world knowledge, 141, 266