CANADIAN THESES

THÈSES CANADIENNES

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

Canada

A THEORY OF STRATIFIED MEANING REPRESENTATION

FOR NATURAL LANGUAGE

by

Tomasz Strzalkowski

M.Sc., Warsaw University, 1981

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in the School of

Computing Science

© Tomasz Strzalkowski 1986

SIMON FRASER UNIVERSITY

July 1986

Approval

Name: Tomasz Strzalkowski

Degree: Ph.D. Computing Science

Title of thesis: A Theory of Stratified Meaning Representation for Natural Language.

Examining Committee:

Dr. Arthur L. Liestman
Chairman

Dr. James P. Delgrande

Dr. Nick J. Cercone
Senior Supervisor

Dr. R. E. Jennings
Department of Philosophy

Dr. Veronica Dahl

Dr. Robert F. Hadley

Dr. Thomas W. Calvert
Vice-President
Research and Information Services

Dr. Graeme Hirst
External Examiner
Department of Computer Science
University of Toronto

Date Approved: _11 July 1986_

Title of Thesis/Project/Extended Essay

A THEORY OF STRATIFIED MEANING REPRESENTATION

FOR NATURAL LANGUAGE

Author:

(signature)

TOMASZ STRZALKOWSKI

(name)

8 August 1986

(date)

ABSTRACT

We introduce a computationally oriented theory of natural language understanding that integrates various levels of language processing with methods of modelling the language denotational base. This theory attempts to bridge the gap between the formal theories of language and meaning. such as "possible worlds" theory and the theory of situations. and artificial intelligence practice. We develop the concept of the Stratified Model as a major processing medium which provides an interface between linguistic input. the "real" universe. and the knowledge base of some hypothetic. intelligent individual.

The major work reported in this dissertation focusses on selected aspects of the Stratified Model which have been identified with the domains of three language transformations and the meaning representation levels they produce. The first of these transformations represents the syntactic analysis of morphologically disambiguated utterances with the categorial grammar CAT. As a result. a set of possible discourses is generated in which each sentence is considered independently of the rest of discourse. Possible discourses are subject to another transformation which integrates them into partially connected and locally coherent discourse prototypes. At this stage. various inter-sentential dependencies. are evaluated. including anaphoric in-text relations. A small number of possible discourse representations which display global coherence are selected from among discourse prototypes and delivered onto the ultimate meaning representation level by the final transformation. From this level a mapping is attempted onto a corresponding universe model. A sequence of further transformations would then extend this mapping. to a real-world interpretation of discourse.

The problems of modelling the language denotational base are addressed within the Theory of Names and Descriptions which constitutes a part of our framework. A layered model of the universe is proposed to capture non-singular interpretations of certain linguistic entities often referred to as functional. generic. mass. or intensional. The uniformity and elegance of our approach is contrasted with the partial and incomplete explanation of this phenomenon given in other relevant research.

For Bozena

Acknowledgements

My thanks and appreciation go out to the following people and places that have seen me through the past three years:

To Dr. Nick Cercone, my Senior Supervisor, for his time and thought that he has given to me, and for his constant encouragement without which this work would not be possible.

To the members of my Supervisory Commitee: Dr. Veronica Dahl. Dr. Tom Calvert. Dr. Ray Jennings. Dr. Jim Delgrande. for their time spent on reading earlier versions of this thesis and suggesting many invaluable comments.

To Prof Ray Jennings for discussions on intensional logics.

To Dr Graeme Hirst of University of Toronto, for being the External Examiner.

To Dr. Bob Hadley for being the Internal External Examiner

To the members of the faculty and the graduate students of the School of Computing Science, for their assistance at various stages of my Ph.D. programme.

To the staff of the School, for their readiness to answer even most obscure of my questions. I reserve here special thanks for Ethel Inglis, our Graduate Secretary.

To the Laboratory for Computer and Communications Research (LCCR) for the use of their facilities. Ed Bryant patiently answered my countless questions regarding TROFF and Scribe. Carol Murchison has been a constant source of moral support and a good friend.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1. Processing Natural Language

Human (or natural) language is the almost universal tool for expressing the variety of kinds of knowledge people can possess. People use their natural language in everyday life without any limitation or effort. They know perfectly well how to use the language, what the language can express, and how to express it. They also understand each other even if they find themselves in novel situations. Clearly people are able to acquire knowledge of the world, store this knowledge somehow, and then manipulate and articulate facts (possibly derived) using natural language solely (or almost always solely). How is it possible for people to understand what they are talking about? In other words, we need to answer the question. What is the nature of natural language that makes communication possible regardless of the differences in native languages used in the world?

In this thesis we develop a new framework for investigating problems of natural language understanding that integrates various levels of language processing with methods of modelling the language denotational base (we shall call it the universe) into a formal, yet intuitive Theory of Stratified Meaning Representation. The theory is formulated from the perspective of the hearer in some hypothetic discourse situation as he attempts to decode the meaning of a message delivered by the speaker. The hearer has at his disposal the Stratified Model, additionally augmented with an individual knowledge base. The knowledge base remains the sole property of its owner, and other individuals are assumed to maintain such knowledge bases as well. The knowledge base contains appropriately encoded information about what its owner knows, believes, doubts, imagines, etc. about the universe. The actual

organization of the knowledge base is immaterial at this point. beyond the fact that it should be con-
veniently accessible for its owner. The Stratified Model provides tools for manipulating information
incoming to. or outgoing from the knowledge base. Information enters (and leaves) the Model through
numerous receptors (or ports) which. for now. we prefer to leave unspecified. Since we are interested
in the linguistic type of data. we shall assume that. at a port. information has the form of discourse in
some source language SL. Perhaps other types of data are processed in a similar fashion. but we shall
ignore these beyond the final appearance in the knowledge base.

The discourse entering a port undergoes numerous transformations before it can be assimilated
into the knowledge base. These transformations should embody the full range of language processing
from lexical and syntactic analysis to advanced semantic and pragmatic evaluation. These transforma-
tions can be performed in parallel or sequentially. but it would be unrealistic. in general. to insist upon
the old-fashioned syntax/semantics/pragmatics distinction. No matter how we eventually organize the
process. however. the resulting representation should accurately reflect the meaning of the original
speaker's communication as perceived by the hearer. This simplified account of natural language
understanding complicates significantly if we allow for conversational discourses where more than one
party makes a contribution. There are many methods for representing the meaning content of natural
language discourse but the one we select must be well-defined. formal. and simple. These are abso-
lutely elementary requirements if our theory is to have a wider significance. Such properties are charac-
teristic for logic and logic-based representations. The Stratified Model employs several different kinds
of representation. We use the categorial language L for representing syntactically pre-processed utter-
ances. and the lambda-categorial language Λ for representing inter-sentential dependencies in
discourse. The final representation of discourse content has not been decided upon yet. but we advo-
cate the form resembling that of *abstract situation*. see Barwise & Perry (1983). In the section 1.2
we discuss a number of logic-based representations used in natural language research and motivate our
choice of L and Λ.

The Stratified Model also provides semantics for verifying the acquired information with respect to the universe. This information is mapped into the universe (interpreted). and the resulting classification is again stored in the knowledge base so that future verification could be directed into the latter. whenever possible. The actual mapping into the universe is not a straightforward process. however. To perform this mapping efficiently an individual has to properly grasp the structure and organization of the universe. We assume that an intelligent individual is capable of deriving appropriate models of the universe. and that he directs the semantic mapping into these models rather than into the original universe. We may call the latter "the reality" but we must remember that the models are real as well. A model is not some abstract representation of the universe. the property that could be attributed to the contents of the knowledge base. A universe model is this same universe but somehow specifically perceived to separate its various aspects objects. relationships. situations. etc. The problem of modelling the universe is an important aspect of language understanding. Although we may assume that the perception of the physical reality does not differ much among people. and subsequently we may consider that the universe is what we actually perceive. the physical world is not all we perceive and model. The social and cultural superstructure of the physical reality is also a subject to classification. and this aspect of the universe modelling will quite likely differ among people. Nevertheless. our approach includes the modelling of the physical reality as well. After all. we may leave room for extra-terrestials.

It has often been assumed that there exists some universal semantics that can classify all linguistic (and other) information depending only on the actual state of the universe. see. for example Lewis (1976). Unfortunately, this kind of semantic classification is not always available for an individual because the individual's perception of the universe is subjective. and this fact is further reflected by the contents of his knowledge base. In this sense. different individuals "run" different instances of the Stratified Model. primarily with respect to different universes they recognize. but possibly with respects to other factors as well. The Stratified Model we construct may not belong to anybody. and we do not make claims about whether or not it has any psychological significance. We shall assume.

however. that our Model is owned by some hypothetical individual who patiently participates in the various discourse situations we analyze in this research. The primary role of the Model is to provide insight into the way an intelligent individual may come to understand natural language utterances. and prepare for a computer implementation of this process.

We do not construct a complete Stratified Model in this thesis. Our main effort is concentrated on selected problems of discourse analysis. We investigate a number of questions concerning discourse cohesion and derivation of a formal representation of discourse content. In the area of text cohesion we propose a general scheme for computing inter-sentential dependences created by pronominals and definite descriptions. We then augment our presentation introducing the concept of remote reference as a consequence of our new account for non-singular terms in language. We also address the problem of automated selection of proper cohesive links in discourse by tracing speaker's intentions. We discuss the phenomena of backward and forward references. direct and indirect references. and comment on the so-called lost links in discourse. We outline the problem of discourse coherence. and finding discourse topic. Finally. we summarize computational aspects of our theory.

It was the author s intention from the start to build a theory that would have a significance for research in Artificial Intelligence and Computational Linguistics. The Theory of Stratified Meaning Representation has been developed with consideration for its future computer realization. There are several related theories of natural language understanding available in philosophy and linguistics. such as the theory of possible worlds. or theory of situations. that lack the computational perspective of the Stratified Model. and despite numerous attempts did not find their way as yet into computing sciences. Nevertheless. these theories present the most successful approaches thus far to the problems of natural language understanding. It is not surprising therefore that our present theory derives from philosophy and formal logic. and works of Frege. Tarski. Ajdukiewicz. Quine. Lewis. Cresswell. and Montague have been of predominant influence. We acknowledge the impact of more recent theories of situations. cf. Barwise & Perry (1983). especially in the domain of modelling the language denotational base. Finally. the insights of Artificial Intelligence research into syntactic and semantic processing of

natural language. knowledge representation. and discourse analysis motivated various detailed solutions which we present in this thesis.

1.2. Logic-based Representations in Natural Language Research

The mystery and power of natural language understanding has long inspired theories about its operation. Mankind. unable to account for human use of natural language. has challenged linguists. philosophers of language. cognitive psychologists. and computing scientists (artificial intelligence researchers) to develop machine intelligence for natural language understanding.

Chomsky's (1957) work began an intensive period of study in linguistics. Linguistic research concentrated almost entirely on problems of syntax which. according to Chomsky's Standard Theory and its later revisions (see Radford. 1983 for discussion and comparison) was the key to further analysis. To catch up with semantic aspects of language processing linguists coined the notion of *logical form* which. unfortunately. was very misleading and had little to do with logic. It soon became apparent that the study of semantics and pragmatics held more promise for answering the questions postulated than a purely syntactic approach. It was quite natural therefore. that researchers turned to logic for help since logic was designed to study semantics.

Logical approaches to semantics proved somewhat successful for dealing with natural languages and the most popular form of logical systems. the first-order predicate calculus (or first order logic [FOL]) became widely recognized and accepted for this purpose. The most attractive properties of FOL. i.e.. elegance of representation. rich semantics. and straightforward inference methods. found their way into computing science research. particularly into the logic programming. Colmerauer (1978). Warren & Pereira (1982).

Shortly thereafter. however. FOL began to reveal some limitations for representing natural language sentence meaning. The limitations derive from the artificial constraints imposed upon the formalism. primarily because of its first-order principle (i.e only single level predication is allowed). It

became painfully evident that notions like time and context could not be expressed in FOL without very artificial and implausible modifications. even when we allow some second or higher-order extensions. Upon that evidence researchers turned to other methods which shared first-order logic's relative computational tractability and would allow for the accommodation of more sophisticated features of natural language semantics and pragmatics. This movement resulted in a number of more or less successful quasi-formal notational systems. empirically tested. which ranged from only slight modifications to the first-order logic. see Woods (1972). Burton (1976). Fillmore (1968). Bolc & Strzalkowski (1982). through knowledge representation languages. for example Bobrow & Winograd (1977). to very advanced methods of graphical representation known under the common term of semantic nets. cf. Schank (1972). Wilks (1973). Schubert (1976). Schubert. Goebel & Cercone (1979). The most prominent and influential among these concepts were *frames*. Minsky (1975). *scripts*. Schank (1975). and *semantic networks*. Quillian (1968). which later evolved into more comprehensive *knowledge representation languages* like KRL. Bobrow & Winograd (1977). FRL. Roberts & Goldstein (1977). NETL. Fahlman (1979). KL-ONE. Brahman (1978). and more recently Krypton. Brachman et al. (1983)

Frames have been designed to describe selected *chunks* of knowledge at arbitrarily chosen levels of detail. In this sense. frames resembled *abstract situations*. Barwise & Perry (1983). but their static and rigid organization became a problem for modelling human knowledge bases. Nevertheless. frames have been used quite widely in AI research. for example to trace focus in discourse. see Sidner. (1979).

Unlike frames or scripts. semantic networks emphasized structural aspects of knowledge. They were significantly more flexible than frames. additionally displaying a close kinship to predicate logic. Although early network systems had less expressive power than predicate calculus. later enhancements enriched the representation with logical operators. quantifiers. and higher-order predications. Schubert (1976). If we disregard methodological considerations. however. we can quickly realize that semantic networks do not offer anything over traditional linear notational systems. Also. the lack of

formal mathematical foundations of the net semantics appeared to be an obstacle for developing more general theories of knowledge representation. A similar criticism applies to knowledge representation languages, which despite efforts to develop a universal notational system, remained mostly application dependent.

Philosophical works of Montague (1974), especially his famous *Proper Treatment of Quantification in Ordinary English* (PTQ henceforth), revived the formal logic approach to natural language understanding. However, because of the complexity and relative computational intractability of some key notions of Montague's theory (possible worlds, intensions, etc.), the approach mostly remained outside the mainstream research in Artificial Intelligence. The ideas expressed by Montague, nevertheless established a new frontier for studying meaning and meaning representations, see Cresswell (1973), Partee (1976), Thomason (1976), Dowty (1976), Heim (1982). Some effort has been devoted to develop an intermediate notational system between first order logic based methods and intensional logic, see for example Webber & Reiter (1977), Moore (1981), Schubert (1982), and Scha (1983). These approaches, however, accommodated only few selected features of Montague's system (for example the lambda operator, higher order predications, temporal operators) often standing in relative disagreement to others. Yet, as we believe, the intensional logic semantics gives the most significant contribution to the theory of meaning thus far, and PTQ shows that natural language might be reduced to a formal notational system which preserves the meaning of original utterances independently of any particular "domain of discourse"

Any meaning representation theory should be formulated at a comfortably abstract level, and its primary role should be regarded as approximating the source language, its semantics and pragmatics, and to some degree, its syntax. The foundations for such research has already been laid by philosophers Frege (1892), Carnap (1947), Kripke (1970), and logicians Ajdukiewicz (1935), Tarski (1936), and others, but a formal notational system was required as a meaning representation medium. Such a tool already existed in mathematics under the name of lambda-calculus. Lambda-calculus offered advantages over classical logical systems allowing for representing non-propositional information.

Elements of lambda-calculus found their way to computer science and Artificial Intelligence (AI). The programming language LISP, considered as the main computer language for AI applications, is partially based on principles underlying lambda-calculus. In the domain of meaning representation the so-called procedural approach used, often implicitly, notions of lambda-abstraction and lambda-reduction.

An alternative for the Fregean-type semantics for natural language has been designated for some time as the theory of situations, see Partee (1972) and Barwise & Perry (1983). The formal logic approach to meaning representation has been challenged once again, this time with the intuitive concept of a *situation*. Unfortunately, the theory does not address computational concerns, and there is little evidence that a situation-based natural language system could emerge in near future. Additionally, insufficient formal foundations of the theory cooled the initial enthusiasm of the AI community. Nevertheless, the theory of situations offers many interesting insights into the problem of knowledge representation that we address in this dissertation.

1.3. Introduction to the rest of the thesis

In Chapter 2 we give an in-depth presentation of the theoretical foundations that prepare us for further discussion. Various aspects of the possible world theory, and its impact on problems of natural language processing are analyzed in section 2.1. We discuss categorial languages, lambda-categorial languages, and intensional logic, and conclude with a brief presentation of Montague's PTQ system, Montague (1974). We also point out difficulties with computer realization of this approach. Situation Semantics, as presented by Barwise & Perry (1983), are analyzed in section 2.2. Numerous similarities between this theory and the possible world semantics are illustrated. We also indicate several problems faced by the situational approach to natural language semantics.

In Chapter 3 we define the notion of Stratified Model, and show how the two major theories of language relate to our framework. Clear distinction is being made between linguistic information processing and the problem of modelling the language denotational base. We also discuss the problem of

ambiguity arising during the information processing. Finally. we explicate the extent to which this thesis contributes to the Stratified Model.

The problem of inter-sentential dependencies in discourse is examined in Chapter 4. We introduce a new framework for treating definite descriptions and pronominals in various contextual situations. Using the formal notions of *context-setting sentence* S_1. and *current utterance* S_2. we distinguish a number of different context environments characteristic for referential and non-referential readings of certain nominal constructions. These include *perfect contexts. imperfect contexts. attitude report contexts.* contexts with proper names. and *conditional contexts.* For each of these cases we formulate translation rules which transform appropriate fragments of discourse into a formal representation in the lambda-categorial language Λ. The rules (numbered 1 to 10) are then accommodated into the recursive formula of a more general transformation F_{n-1} within the Stratified Model. We show how this transformation can be employed to process backward. forward and indirect reference cases.

In Chapter 5 we continue the discussion of inter-sentential dependencies. and we extend our investigation beyond singular terms of language. A new treatment for non-singular noun phrases is proposed in the Theory of Names and Descriptions (TND). The theory suggests a layered organization of the universe and defines the notion of relative singularity of certain types of objects as equivalence class for the naming level membership relation. We introduce the notions of *superobject* and *coordinate* to uniformly represent various kinds of compound concepts. including generic objects. mass objects. and intensional objects. TND shows that the philosophical notion of intension as a function over possible worlds is not indispensable. and thus offers renewed hope for a computational treatment of intensionality. The theory adds two remote reference rules (Rules 11 and 12) to the transformation F_{n-1}. These rules translate cohesive links relating linguistic terms referring to objects classified into different naming levels. Finally. we define the notion of *discourse prototype* as the representation of discourse derived by the transformation F_{n-1}.

In Chapter 6 we turn our attention to the general problems of discourse analysis: cohesion, coherence, and maintaining a discourse model. We sketch the definition of the transformation F_n in the Stratified Model that takes discourse prototypes and produces ultimate discourse representations. A number of factors that influence the selection of proper cohesive links between parts of discourse are described and classified according to their deterministic force. These factors range from purely lexical number/gender agreement for pronominals and their antecedents to discourse coherence factors that trace the speaker's intension. We address the notions of *active context* (focus), *discourse topic*, and *topic entity*, and examine their roles in the antecedent taking process. We also analyze the phenomenon of repairing effects of incorrect presuppositions by revising the discourse model maintained by the hearer with the *backward adjustment process*. We suggest methods for recovering *lost links* and unobserved forward references in discourse.

Chapter 7 is devoted to computational aspects of the transformations discussed in Chapters 4 to 6, as well as the Stratified Model in general. We argue that our approach promotes future computer realization, and we outline a possible way to implement various parts of the Stratified Model. We also point out a number of difficulties that may arise.

Future directions for this research are discussed in Chapter 8. We propose numerous ways to expand the present version of the Theory of Stratified Meaning Representation. In several places we suggest novel concepts which have yet to be examined. Among them, the extension to the transformation F_{n-1} by *creative contexts* is considered. Finally, we present a summary of the major contributions of this thesis to natural language understanding research.

Chapter 2

Advanced Semantic Theories

2.1. Possible World Semantics

The term *possible world semantics* does not refer to any consistent theory understood as a more or less formal system with clearly defined set of primitives and rules. Instead, it is taken as the common label for the mainstream research in linguistics, philosophy, and logics initiated a century ago by German logician and philosopher Gottlob Frege. In hands of such theorists like Ajdukiewicz, Tarski, Lukasiewicz, Carnap, Quine, Cresswell, Lewis, Thomason, Kripke, Montague, and many others Fregean concepts have been developed to the most successful approach thus far to problems of representing meaning of natural language expressions. All these researchers, however, viewed the problems from different perspectives, with different goals in mind, and they often formulated their findings at different dimensions. What have these efforts had in common, except pursuing, as any philosophy, a more satisfactory answer for the general problem of meaning? You may say, possible worlds are what distinguishes these workings from others, but it won't be quite accurate. Many authors do not seriously mention this term speaking of points of reference, indices, contexts, or whatever. What connects all these concepts is their common predecessor, the Fregean notion of **sense**. There is no commonly accepted definition of a possible world, although there were (unsuccessful) attempts to formalize this notion, for example by identifying possible worlds with Tarskian models; see Montague (1974c) for discussion. In this section we shall try to systematize this notion along the lines proposed by Lewis (1976), Cresswell (1973), and Montague (1974a-f). In the same sense our use of the term *possible world semantics* will mostly refer to the Montague approach to semantics of natural

language. commonly known as *Montague Semantics*, and, in a less degree, to the works by Cresswell, Lewis and Quine. As our main concern in this thesis is to suggest a general strategy for building computationally oriented natural language understanding systems, we are also particularly interested in ways in which the possible world semantics relates to the problems of natural language processing at some pre-semantic levels. The most explicit exposition of such a "syntactic connection" can be found in Ajdukiewicz (1935). and then Lewis (1976) and Montague (1974ef). Montague's *Proper Treatment of Quantification on Ordinary English* (PTQ henceforth) will be the most often referred to system. For that reason we include in this section a presentation of the categorial approach to English syntax. which became tightly connected with Montague treatment of natural language.

The description which follows is based on the work of Cresswell (1973). Montague (1974a-f). Lewis (1976). Thomason (1976). Partee (1976). Kripke (1970). and others. and on my own work. Since all authors attacked the same problem from slightly different positions. since they stressed different aspects of the problem and finally since they used unique notations. I have attempted to unify my descriptions of their work and standardize on a single notation.

2.1.1. Categorial Grammars

The notion of syntactic category in formal languages is widely recognized to be due to Ajdukiewicz (1935). who initiated systematic research on the connection between syntax and semantics. Frege (1892) provided the philosophical motivation for research in this direction with what is now generally known as the **Frege Principle** (the meaning of a language expression is a function of the meanings of its parts). The Frege Principle has been stated in connection with natural languages and it still has the character of a hypothesis.

Categorial languages have been invented to test the hypothesis. Whether natural languages are categorial languages remains an open question. The recent work of Montague (1974d-f). Lewis (1976). Partee (1976). Cresswell (1973) and others put forward strong arguments that this may. in fact. be

the case. Although their results are promising and intriguing, they remain fragmentary and little progress has been made since Montague's original papers were published over a decade ago. The current trends in linguistics and the philosophy of language indicate that the categorial approach lacks universal recognition as the "correct" paradigm for natural language understanding. Furthermore, as obstacles begin to accumulate (we shall discuss these later), linguists have recently turned toward more classical methods originally proposed by the transformational school of thought.

Alternatively, the "categorial" approach is one of the few which has attempted to give a satisfactory explanation of natural language semantics and pragmatics. From our observation, the major shortcoming of the categorial approach lies in its neglecting the role of syntax in language analysis. Although some work has been done to incorporate transformational issues in the syntax of categorial languages, Partee (1976), they do not go far beyond the limited scope of Montague's grammars.

We now present a description of categorial languages, their syntax, semantics, and pragmatics, starting from a simple notion of propositional languages and then extending it successively to a general concept of lambda-categorial languages.

2.1.1.1. Propositional Languages

Definition 1 [Cresswell (1973)]

By a propositional language L we mean an ordered pair $<\Delta\ S>$ where $\Delta = <\Delta_0 \quad \Delta_k>$ is a finite sequence of pairwise disjoint, possibly empty, finite sets, and S is the smallest set such that

(i) $\Delta_0 \subseteq S$

(ii)   if $\delta \in \Delta_n$ $(1 \leqslant n \leqslant k)$, and $\alpha_1, \ldots \alpha_n \in S$ then $<\delta, \alpha_1, \ldots \alpha_n> \in S$ $\square$

The definition above, though very general, is intuitively simple. Call $\Delta^* = \bigcup_{i=0}^{k} \Delta_i$ the set of all symbols of L. Members of Δ_0 will be then simple (or basic) symbols, and for any $1 \leqslant n \leqslant k$, $\delta \in \Delta_n$ will be a

n-place functor (function constant) over other members of S. The set S is then the domain of all well formed formulae of L. To give an example let us see how the familiar propositional logic can be described using the definition above. We have

Δ_p = set of all propositional symbols (variables)

$\Delta_1 = \{\neg\}$

$\Delta_2 = \{\rightarrow \cup \,\&\, \equiv\}$

$\Delta_n = \emptyset$ for $n \geq 3$

Definition 1 specifies the syntax of propositional language L. It uses implicitly the notion of syntactic category. Indeed we can classify all expressions in L as belonging either to category of saturated sentences S or to one of categories of unsaturated functors which provided with a proper number of arguments from category S produce new elements of category S. Thus the category S consists of two classes of elements. These being taken from Δ_0 which we shall call basic elements. and those obtained through application of functors - thus called compound elements of category S.

Let us adopt the following notation. We shall say that an n-place functor symbol δ belongs to syntactic category s/(s ... s)$_{n\ time}$ when provided with n elements of category S produces an element of the same category S. We can now regard sets Δ_0 ... Δ_1 in Definition 1 as sets of basic elements of categories s. s/s. ... s/(s ... s)$_{k\ time}$. respectively. Our notion of propositional languages restricts us to the fact that the only compound elements in the language are those of category s. i.e. the elements of S Δ_0. This of course will not be the case in general. when we move to higher-order categorial languages

The last thing worth noting in this context is the linguistic aspect of the notion of syntactic category. If we regard $\Delta = \langle \Delta_0 \ldots \Delta_1 \rangle$ as a lexicon of basic expressions in respective categories. we have a possibility to specify our language L as a set of phrase structure rules.

We have

$s \rightarrow \alpha$ for every $\alpha \in \Delta_0$

$s/(s \ldots s) \rightarrow \delta$ for every $\delta \in \Delta_n$, $1 \leqslant n \leqslant k$, and

$s \rightarrow s/(s \ldots s)$. $s \ldots s$ (i times)

Example Cresswell (1973)

Let L be a propositional language described by the following phrase-structure rules

$s \rightarrow \alpha$

$s \rightarrow \beta$

$s/(s) \rightarrow \gamma$.

$s/(s.s.s) \rightarrow \delta$

then the expression

$<\delta. \alpha. <\gamma. \beta>. <\gamma. <\delta. \alpha. \beta. \alpha>>> \in L$

and can be assigned the phrase marker shown in Figure 2 1

2.1.1.2. Generalized Categorial Languages

We now define a generalized notion of a categorial language based on principles introduced in the previous paragraph The definition has been derived from the ones given by Cresswell (1973) and Lewis (1976)

In the case of propositional languages we had one basic category s (of sentences) and a finite number of derived categories of propositional functors Moreover the only category which was allowed to have non-basic elements was the category of s All this will be remedied presently Prior to the actual definition of categorial languages let us establish the extended notion of syntactic category

Definition 2

Let BCAT be a finite non-empty set of basic categories. In practical applications BCAT has usually no more than 2 or 3 elements Define the set of syntactic categories CAT to be the smallest set such that
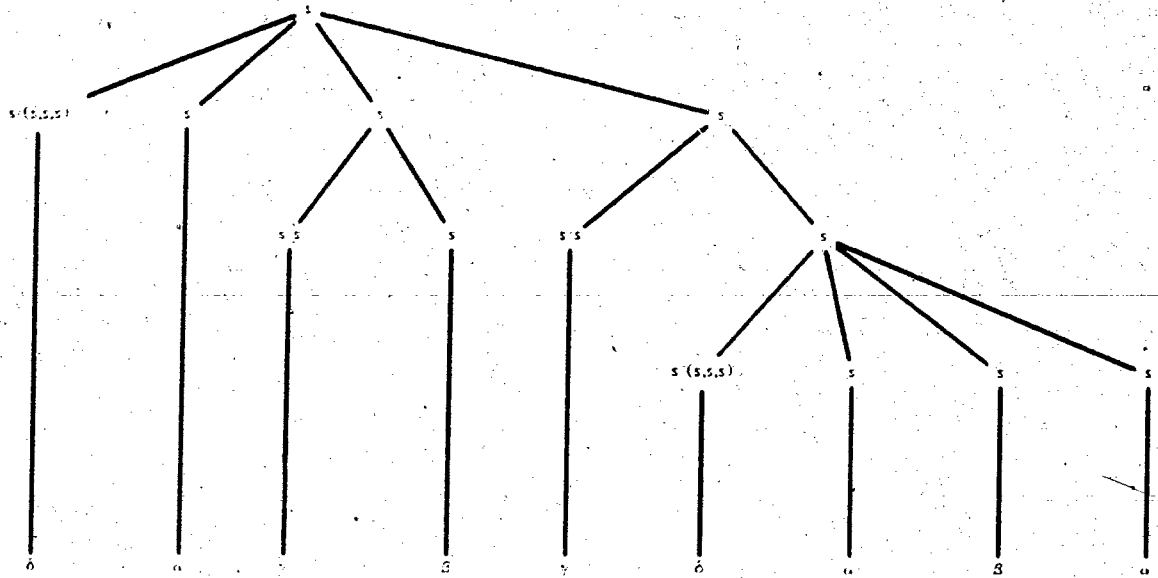
Figure 2.1. Derivation in propositional language.

(i)   BCAT $\subseteq$ CAT

(ii)   if $\tau \ \sigma_1 \quad \sigma_n \in$ CAT then

$<\tau \ \sigma_1 \quad \sigma_n> \in$ CAT $\sqsubset$

We shall call $<\tau. \ \sigma_1. \ \dots \ \sigma_n>$ the derived category of n-place functors such that when provided with n

arguments of categories $\sigma_1 \quad \sigma_n$ respectively (in this order) will produce a member of category τ.

With this concept we are ready to give a formal specification of a generalized categorial language.

Definition 3

By a categorial language L we mean a ordered 4-tuple $<$BCAT. Δ. B. E$>$ where

(i)   BCAT is a finite. non-empty set of basic syntactic categories

(ii)   CAT is the set of all syntactic categories as defined by Definition 2

(iii) $\Delta = <\Delta_{\sigma_1}, \ldots \Delta_{\sigma_i}, \ldots >$ is the sequence of finite, disjoint sets of which all but finitely many are empty. Δ will be called the lexicon of L

(iv)   B is a function from CAT to Δ such that for any $\sigma \in$ CAT. $B(\sigma) = \Delta_\sigma$. $B(\sigma)$ is called the set of basic expressions within category σ.

(v)   E is a function whose domain is CAT and such that

    (a)   for any $\sigma \in$ CAT. $B(\sigma) \subseteq E(\sigma)$.

    (b)   for any $\tau, \sigma_1, \ldots \sigma_n \in$ CAT. if $\alpha_1, \ldots \alpha_n \in E(\sigma_1), \ldots E(\sigma_n)$ respectively, and $\delta \in E(<\tau \, \sigma_1, \ldots \sigma_n>)$ then $<\delta, \alpha_1, \ldots \alpha_n> \in E(\tau)$. $E(\sigma)$ is called the set of all well-formed expressions of category σ. $\square$

The above definition fully establishes the syntax of categorial languages of the sort we shall call, after Cresswell (1973) **pure categorial languages** Later we shall show how the notion can be modified to specify the concept of λ-categorial languages. We now relax our formal discussion to show how pure categorial languages could be applied to natural language analysis. We shall return to the formal environment afterwards, and turn our attention to semantic and pragmatic side of the problem

2.1.1.3. English as a (pure) Categorial Language

To design a pure categorial grammar for English one would need at least three basic syntactic categories These are the categories of sentences, names, and common nouns which we shall denote using S, N. and C symbols. respectively; Lewis (1976). This categorisation can be reduced to two basic categories, but we postpone this discussion to a subsequent section.

Let us consider what kind of expression could be found in these basic categories. Within category N our lexicon is expected to contain proper names like "John". "Mary". etc., while common nouns like "queen". "man". etc., create the lexicon part for category C. Obviously. category S has no basic expressions in it. With the basic categories we can define our first derived category of 1-place intransitive

verbs as S/N. thus obtaining the first context-free phrase-structure rule in the grammar.

(1) S → S/N. N

Ignoring. for the moment. the problem of attaching some transformational component to refine the proper word-order we can now generate simple sentences of English using the above rule. Suppose. therefore. "sleeps" is among the basic expressions of category S/N. According to Definition 3 from previous section. the sentence

(2) <sleeps. John>

is generated in category S

In the next step we would like to accommodate sentences like

(3) Everyone sleeps

in our framework. The quantifier phrase "everyone" cannot be treated as a name nor as a common noun. What we need for such expressions is a category which given an intransitive verb from category S/N can produce an expression of category S. Such a category will be S/(S/N). so we have the next rule in the grammar

(4) S → S/(S/N). S/N

Now observe that common noun phrases like "a queen". "every man". etc. behave (both syntactically and semantically) like quantifier phrases "someone" and "everyone". We want them to place into the same category S/(S/N). Observe also that category of S/(S/N) has no basic expressions (lexemes). This derived category will change. however. when we redefine our categories later on.

Now we are able to include in our fragment of English sentences like:

(5) John sleeps.

     Everyone sleeps.

     A queen sleeps.

From this position it is only a formal operation to establish a derived category for indefinite articles "a".

"every". "some". As they take common nouns to create noun phrases they must fall into category

(S/(S/N))/C. It should be also relatively easy to imagine that the 2-place transitive verbs like "love".

"seek". etc. have to be classified within category (S/N)/(S/(S/N)). i.e. the entities which given an ele-

ment of category S/(S/N) (noun phrase) yield an expression of category S/N (intransitive verb

phrase). Indeed. expressions like

(6)   "loves someone"

       "seeks a queen"

can be regarded as syntactically equivalent to intransitive verbs. Some problems arise in this situation.

however. since we still need transitive verbs in category (S/N)/N to account for expressions like

(7)   "loves Mary".

As Lewis (1976) suggests. one possible solution is to void the basic category N of proper names. and

establish some equivalent notion of **pseudo-names** which will than fall into category S/(S/N). In

this way we reach the desired unity of representation. The operation cannot be properly explained on

purely syntactic ground so we postpone this question to the section on semantics and pragmatics of

categorial languages. What we can say now is that the lexical elements newly appended to category

S/(S/N) are no longer treated as names. They will be regarded as **lexical pseudo-names** which in a

particular pragmatic context can denote names. The category N is voided but not discarded altogether

Its role will become clear when we define meaning postulates for categories.

At this point. how to construe a categorial grammar for English should be clear enough. Before

we give an concrete example. let us refine the categories distribution along the lines suggested by

Montague (1974f). Partee (1976). and Thomason (1976). We abandon the category C of common

nouns. and establish only two basic categories t and e which are equivalent to our last concept of

categories S and N respectively. As Montague (1974f) observed in his work some derived categories

may constitute 2 or more copies. see Delacruz (1976). which although behave differently in purely syn-

tactic environment, have identical semantic significance. Clearly. for some categories A and B we can define not only one derived category A/B but also A//B. A///B. .... All of these categories can be said to take an element of category B to create an element of category A. The difference lies in the way the syntactic process could take place. But, since our main goal is establishing a meaning representation. uncovering such properties of language has a particular significance. Therefore. except for a primary copy. some derived categories can also have other copies in our grammar. We shall regard them. at least for now. as distinct syntactic categories. In our discussion thus far we intentionally avoided specifying how the syntactic derivation should actually be conducted on the surface level to get a well-formed English sentence (with attributes like word order. tense. case. person. etc. specified). We shall still keep this convention. The syntactic derivation problem seems to belong to the domain of a transformational component of the grammar and this is of no interest to us.

We are now in a position to present an example categorial grammar for a fragment of English. According to Definition 3 all we need is to specify basic categories and the lexicon. Table 2.1 lists the most frequently referred to categories of the present fragment, and sets some notational conventions. The example has been derived from Montague (1974f). adapted and reduced to the requirement of the present discussion. The lexicon for our fragment is presented in Table 2.2.

One can safely proceed further to cover an even wider subset of English with a grammar defined along the lines just introduced. But the framework is as yet inadequate. and actually the concept of pure categorial languages is too weak to define an acceptable grammar for English (and I think for any natural language in general). To illustrate this point. let us discuss the problem of quantifier scoping. Lewis (1976) gives the following example.

(8)   Every boy loves some girl.

The sentence is ambiguous in two ways. yet one of the readings evades treatment in a pure categorial grammar. In one sense. where for every boy there exists a possibly different girl such that he loves her. we have the p-marker as shown in Figure 2.2 below. There is, however. another possible reading for

Categories in the fragment of English		
name	abbrev.	nearest linguistic equivalent
t		proposition, declarative sentence
e		proper name, entity
t/e	IV	intransitive verb phrase
t/IV	T	term, noun phrase
IV/T	TV	transitive verb phrase
IV/IV	IAV	IV modifying adverb
t//e	CN	common noun – –
t/t		sentence modifying adverb
IAV/T		IAV-making preposition
IV/t		sentence taking verb phrase
IV//IV		IV taking verb phrase

Table 2.1. Categories defined for the present fragment of English

$B(t) = \varnothing$
$B(e) = \varnothing$
$B(IV) = \{run, sleep, walk \quad \}$
$B(T) = \{John, Mary \quad \}$
$B(TV) = \{find, love, seek....\}$
$B(IAV) = \{rapidly, slowly....\}$
$B(CN) = \{man, woman, queen \quad \}$
$B(t/t) = \{necessarily....\}$
$B(IAV/T) = \{in, about....\}$
$B(IV/t) = \{believe that, assert that, .\}$
$B(IV//IV) = \{try to, wish to....\}$

Table 2.2. The lexicon for the present fragment
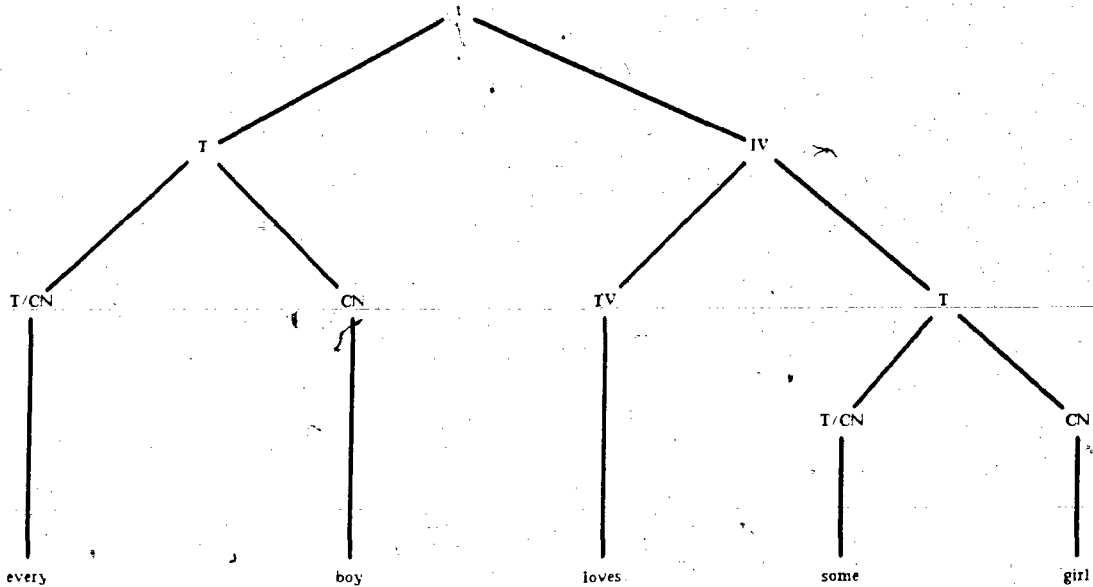
Figure 2.2. Derivation in pure categorial grammar.

the sentence above which comes as

(9). There is a girl such that every boy loves her.

We need to recognize the expressions which contain variables, and what is more important, to assign them to a category. For example. we need a rule in our grammar which would account for the derivation shown in (10).

(10) every boy loves some girl → <(every boy loves x). some girl>

Two questions arise. We need to explain the role of the variable x in our grammar, and to find a way to form a predicate out of "every boy loves x" to make the above derivation consistent with the Definition 3 of categorial languages from previous section.

The solution for the first of these problems comes easily when we realize that "every boy loves x" is a sentence. thus falls into category t. and that x plays the role of direct object for "loves". therefore must be of category $T = t/(t/e)$. It follows that we must extend our lexicon of basic expressions for

category T to also contain an infinite number of individual variables. (or rather linguistic pseudo-variables) so we have

(11) $B(T) = \{$John, Mary,... $x_0, x_1,...\}$

The answer to the second question does not seem to be in scope of pure categorial languages. One way to accomplish that is to employ the concept of **abstraction** most widely known as λ-abstraction. The expression

(12) $(\lambda x \text{ (every boy loves } x))$

is actually 1-place predicate, where λx can be read as "an x such that". Now, the requirements of Definition 3 can be fulfilled. The derivation (10) is refined as follows:

(13) every boy loves some girl $\rightarrow$ $<(\lambda x \text{ (every boy loves } x))$, some girl $>$

Any λ-expression of the form $(\lambda x (Q x))$ behaves exactly like an intransitive verb. it must then fall into category IV. † One could observe that we can use a similar operation to represent the weak reading of the sentence in question (Figure 2.2). Indeed. the λ-abstractions can be used almost without limitation and are very a powerful representational device. However. we need to extend our definition of pure categorial languages to include the notion of λ-abstraction somehow in it. We shall return to this problem later on.

We now turn our discussion (for the first time) to the central issue of this proposal: the semantic and pragmatic aspects of the formalism introduced. The following section discusses the problems at length.

2.1.2. Semantics of Pure Categorial Languages. Frege Principle

When we talk about semantics of a language or a syntactic system we ask what is the correspondence between expressions of the language and a world in which the language is intended to

---

† Since $(\lambda x (Q x))$ is from category IV $= t/e$. and $(Q x)$ falls into category t. the abstractor λx goes into the new derived category of $(t/e)/t$. i.e. the category of expressions which when applied to an element of category t yield an element of category t/e.

be used. In other words we want to know what is the **meaning** of a particular syntactically well-formed expression of the language, again in the context of a particular use.

The most straightforward way to do so is to associate with our language L a domain, say D, of "things" of which the language is to communicate, and a function, say V, which would fix the correspondence between the language and its domain D. Let's call the function V the interpretation of L in D. We would like V to have the following, very general property:

(1)   for any $\alpha \in L$, $V(\alpha) \in D$

The definition is indeed very general, and intuitively insufficient for determining meanings of categorial languages. Even in the case of propositional languages we have at least two different classes of expressions within a language which, we feel, need different interpretations. Cresswell (1973) observes that the interpretation system for a propositional language L is an ordered pair $<V, D>$ such that

(2)   if $\alpha \in \Delta_0$ then $V(\alpha) \in D$.

(3)   if $\delta \in \Delta_n$ then $V(\delta) \in D^{D^n}$

i.e. interpretations for basic expressions in the only basic category s (with the lexicon Δ_0) are "things" from domain D, while the interpretation for an n-place functor from category $s/(s, \cdots, s)_{n-times}$ is an n-place function from D into D. Cresswell defines then the closure of operation V as V^- thus obtaining a full account of interpreting expressions in propositional languages.

(4)   if $\alpha \in \Delta_0$ then $V^+(\alpha) = V(\alpha)$.

(5)   if $\delta \in \Delta_n$ and $\alpha_1, \ldots, \alpha_n \in S$ then

$$V^+(<\delta, \alpha_1, \ldots, \alpha_n>) = V(\delta)(V^+(\alpha_1), \ldots, V^+(\alpha_n))$$

If we suppose that V (and V^+) expresses the meaning of an expression of L, the rules above reflect the important principle, widely known as **Frege's Principle**, which says that the meaning of a compound expression can be obtained as a function of meanings of its constituents. Yet we do not know what the meaning itself is. For a propositional language in which only well-formed expressions are

sentences. we can say that the meaning of a sentence (basic or compound) will be its truth or falsity in D. In this sense, one can introduce (after Cresswell (1973)) the concept of model for propositional languages which would be an ordered triple $<D, T, V>$ such that $T \subseteq D$, and $\sigma \in S$ is **true** in the model if and only if $V^+(\alpha) \in T$. What then is the domain D?

As Cresswell (1973) suggests. let D be a set of **propositions**, i.e. abstract entities which somehow realize meanings of sentences. It follows then that propositions may be true or false, those of T being true for a particular model or world. One can then easily imagine another world in which the set T of true propositions would be different than this of some "real" world. Thus we are arriving at a key notion in studying the meaning of sentences i.e. to the concept of possible worlds. A **possible world** is a "state of affairs" which could possibly happen. i.e. no possible world should contain the proposition that $\alpha \,\&\, \neg\alpha$ is true. One can go further and say that if we can have possible worlds, we could have impossible worlds as well in which such expressions as $\alpha \,\&\, \neg\alpha$ or "square circle" are recognized as truths. In this light we can identify propositions with sets of worlds (possible and impossible) in which they are true. One can agree that two propositions are logically equivalent if they have the same set of *possible* worlds in which they are true. Yet. the equivalence is not enough to say that both mean the same. For example the sentences "2 + 2 = 4" and "7 is prime" are realized by propositions which are true in all possible worlds (disregarding the digital system conventions and setting the decimal system as the only one). yet they mean different things. On the other hand, any two propositions are identical when their equivalence can be extended to the impossible worlds as well. In this sense. as Cresswell (1973) suggests. the sentences α and $\neg\neg\alpha$ are **not** identical since the interpretation of negation can change unexpectedly in some impossible world. But we still do not know when two sentences **mean** the same and what would give us an account of what is the meaning.

One proposal to account for meaning comes from Lewis (1976) and Montague (1974a-f) but before we present their point of view we complete and broaden our definition of semantics of categorial languages. Recall that in the case of pure categorial languages we had, in general. a set BCAT of basic syntactic categories (instead of merely one category as in propositional languages). and the

corresponding set CAT of categories described by Definition 2. We can now say, following Cresswell, that a model structure for a pure categorial language L be an ordered pair $<D, T>$ such that D is a function from CAT which establishes domains for categories, i.e.

(6) if $\sigma = <\tau, \sigma_1, ..., \sigma_n>$ then $D(\sigma)$ is a set of (total or partial) functions from $D(\sigma_1) \times ... \times D(\sigma_n)$ into $D(\tau)$.

In this way, $D(t)$ is the set of propositions (the domain of category 0 or t), $T \subseteq D(t)$, $D(e)$ is the set of individuals (i.e. the domain of category 1 or e), while $D(\tau/\sigma_1....\sigma_n)$ be a set of all n-place functions from domains of categories $\sigma_1, ... \sigma_n$ into the domain of category τ. Having established the domains for categories we can give a value assignment coordinate to our model structure to get a model for a pure categorial language L.

Definition 4

A model for a pure categorial language L will be an ordered triple $<D, T, V>$ such that D and T are as before, and

(a) for any $\sigma \in$ CAT, if $\delta \in B(\sigma)$ then $V^-(\delta) = V(\delta) \in D(\sigma)$.

(b) if $\tau/\sigma_1 ... \sigma_n \in$ CAT then $V^-(<\delta, \alpha_1, ... \alpha_n>) = V^+(\delta)(V^+(\alpha_1), ... V^+(\alpha_n))$ where $\delta \in E(\tau)$, $\alpha_i \in E(\sigma_i)$, $1 \leqslant i \leqslant n$, and V^- is the closure of V as defined above. $\square$

Now if we say that the function V expresses the meaning of an expression of L in a given model we can easily observe that pure categorial languages obey Frege's Principle. One could restate this property saying that in any model $<D, T, V>$ for L there is a unique value of V^+ for every expression of L. This principle can be posed in an even stronger version, see Cresswell (1973), but it seems to be irrelevant here. It would be not difficult to see that the above statement won't work for natural languages, even for subsets which can be described in terms of pure categorial languages. Let us borrow from Cresswell a simple example which illustrates this point. Take the sentence, quoted here from Cresswell (1973):

(7)   Women lay observers at Vatican Council.

The ambiguity lies in the word "lay" as being assigned to the category of verbs or adjectives, taking aside the lexical ambiguity of the verb "lay". What seems to be the case in this context is the fact that the linear structure of natural language expressions is insufficient to be semantically analyzable. Rather what we need is the expression phrase-marker (p-marker) which exhibits the internal, categorial structure of the expression. From this point on when talking of a sentence meaning or sense we shall often consider the sentence as its p-marker in some categorial grammar.

The question naturally arises concerning what kind of "things" one could expect in the domain of basic categories. In other words, if we name basic categories of a pure categorial language by successive integers, what is the nature of D_i $(0 \leqslant i \leqslant n)$. We have already seen that for category 0 of sentences D_0 is the set of propositions, and we said later that they may be identified with sets of possible worlds in which they are true. Propositions are therefore functions from possible worlds into truth values. Propositions then can be regarded as **generalized truth values** which, depending on a possible world, become truth or falsity. Generalizing the notion even further, we ultimately arrive to the following.

Let $i \in$ BCAT be any basic category of some pure categorial language L. The domain D_i for that category will consist of a finite or infinite collection of **generalized objects** which, depending on a possible world, may or may not realize some sort of "things" in that world, and which are intuitively relevant to the expressions of category i. Therefore a domain D_i contains functions from possible worlds to subsets of these worlds. We shall call the functions **intensions** of expressions from category i. Values of intensions in a particular possible world will be called **extensions** of expressions from category i. Thus the domain D_i for a basic category i will be the set of intensions of expressions from i into the set of their extensions. It is not difficult to observe that some elements from i may have empty extensions in some possible worlds.

At this very general level the notions of intension and extension as defined above, seem to be consistent (or at least reducible) to the similar notions introduced by Cresswell (1973), Lewis (1976) and Montague (1974d-f). It was necessary, however, to ascend to this higher level to see that they coincide. To be more specific, let us return for a moment to the sample categorial grammar for English introduced in the previous section. Recall that we had agreed to recognize two basic categories in English. They are t - the category of sentences, and e - the category of names. Extensions of elements from category t are of course truth values. Extensions of elements from category e are the "things" named by them. As we said before, the category e has no lexical elements i.e. $B(e) = E(e) = \emptyset$. It contains however non-lexical elements, namely "traces" of those elements we have moved to the category $t/(t/e) = T$.

A lengthy discussion of this very problem can be found in Kripke (1972). Recall that in section 2.1.1 we moved the names in our fragment of English from category e into category $t/(t/e)$. In this way proper names of language have been identified with their definite descriptions. Kripke (1972) call names the **rigid designators** i.e. such that in any possible world they designate the same things (unless they do not exist in these worlds). Certainly definite descriptions of names are not rigid designators. If we say "N = x such that $\Phi(x)$" we mean that in a certain possible world it is the case (it could be) that (contingently) N = x. But in another possible world, x can change its properties or can even fail to exist. This does not mean that N does not exist. It does mean that in this world designators of N and x are different, or even that in this world N is no longer called N (perhaps x is called that). This does not imply that N is not N in some possible world. N can be still called N, called different (contingently), or ultimately not called at all. Kripke gives an example in which the name "meter" is designated in one possible world (our real world) by the length of some metal stick deposited in Paris.† Yet, one can easily imagine that in another possible world people could choose a different object as a length pattern, still calling it "meter". But our "real" meter would not cease to be a meter even if it would no longer be recognized as such (it would have length of 7.385497 m in that

---

† Presently, a more accurate length pattern is accepted in metric system.

world. for example) Some doubts may arise when one considers the following example

(8)   X is the man born from Y and Z

What is therefore the designation of X if in some possible world Y and Z had no children? If we assume "existence" as a property (contingent) then clearly. in the world in question X lacks this property

The discussion above explains the necessity of keeping category e as semantically important. and sheds some light on how the intension of an expression of this category works

One can now clearly see that Cresswell's meanings i e function V are what we have just called intensions. at least for basic categories. But the earlier given definition of V over derived categories leaves no doubts that what Cresswell takes as meanings are actually intensions. Let us now turn to the Lewis (1976) paper and look at what are the intensions of derived categories. According to Lewis (1976)

> An intension for a [category] c/c_1 ...c_n where c. c_1 ...c_n are any categories basic. or derived. is any n place function from c_1-intensions. .... and c_n-intensions to c-intensions

To give an example let us consider the derived category t/e of intransitive verb phrases. An intension of an element $\alpha \in$ t/e or $Int(\alpha)$. or t/e-intension. is therefore a function from e intensions to t intensions. That is. as Lewis (1976) says

> it is a function from functions from indices [or as we call it so far. possible worlds (T S)] to things to functions from indices to truth values

The intension of "sleeps". for instance. is the function Int_{sleep}. such that given any generalized name i.e.. any function Int_1 from indices to things. yields another function Int_2 from indices to truth values such that

$$Int_2(i) = \begin{vmatrix} TRUTH & \text{if } Int_1(i) \text{ is something which sleeps at } i \\ FALSITY & \text{otherwise} \end{vmatrix}$$

Similarly. one can explain how intensions of more complex categories are built. and though it may become more and more complicated as the complexity of a category grows. it gives us simple method

to define what Cresswell called meanings.

Yet Lewis (1976) is not happy with this. He claims that intensions do much of what meaning does but they are not meanings yet. In his view, the meaning of an language expression must also mirror the expression p-structure in the grammar. That is, the meaning of an expression of category $c/c_1 \ldots c_n$ is the tree having at its top node the pair $<c, Int_a>$ where $\alpha \in E(c)$, and immediately beneath are n nodes $<c_1, Int_{a_1}> \ldots <c_n, Int_{a_n}>$ (in this order) where $\alpha_i \in E(c_i)$. If any of categories c_i is derived then the tree continues until all terminal nodes contain basic categories. To support this thesis Lewis (1976) gives two example English expressions which, having the same intensions, differ finely in meaning. They are quoted here from Lewis (1976):

(9)   Snow is white or it is not

(10) Grass is green or it is not.

Certainly both sentences are tautologies and have the same intension (i.e extensions in all possible worlds) Yet they differ in meanings as their components "Show is white" and "Grass is green" have obviously different intensions

We leave this discussion until we have introduced the generalized concept of categorial languages known as λ-categorial languages. We'll try then to refine some notions introduced here as well as restate the definition of intension and extension to better suit our final purpose. For now the reader should have enough material to grasp the idea of semantics for categorial languages in which two notions have a key character. First, the meaning for expressions in a categorial language (whatever it finally is) is derived based on Frege's Principle. Second, the notion of possible worlds has allowed flexibility over the strictly formal model-theoretic approach to deal with semantic complexity of the scale comparable to that of natural languages.

Before we turn to other topics, let's briefly discuss the problem of possible worlds, which has not been yet determined beyond some intuitive hints. It should be relatively clear by now that the concept of possible worlds is used to create contexts or a pragmatic environment in which language

expressions are assigned meanings. We use the more neutral term and call these contexts **points of reference** or **indices** Lewis (1976 Montague (1974d-f). The most explicit presentation on what such an index can be found in Lewis (1976). We follow his suggestions

Let us take indices as ordered n-tuples of various coordinates upon specification of which we can determine extensions of various kinds of language expressions (whenever extensions exist) † First and the most important coordinate of an index is that of possible world, where a possible world is thought of as a particular "state of affairs" as we intuitively assumed thus far. Yet there seem to be other coordinates which must be taken into account to determine language object extensions. Several contextual coordinates will emphasize some parts of a possible world which are of greatest interest to us. The **time coordinate** will help us to evaluate sentences like "Today is Thursday": the **place coordinate** will be crucial with sentences like "There are many people here" The **speaker coordinate**, **audience coordinate** and **indicated-object coordinate** are necessary to determine extensions of such pronouns like "I" "you" "he" "that boy" etc. Lewis mentions even of the **previous-discourse coordinate** for treating sentences like "The aforementioned pig is Porky". Finally, Lewis discusses the **assignment coordinate** which will give the values to any free variables that occur in language expres-sions. In this view, the assignment coordinate is an infinite sequence of "things" Each variable employed in the language will be then a name ‡ having as its intension (for some number n) the **nth variable intension** i.e. a function from indices to **nth variable extensions** which in turn are nth terms of assignment coordinate at indices

One could imagine some further coordinates to index, and Lewis (1976) indeed proposes more At this stage further details may be relevant only at a very practical level The next two sections will complete our theoretical consideration of categorial languages by discussing syntax and semantics of λ-categorial languages. In the last section of this chapter we shall revisit our fragment of English as

---

† Some language objects won't have extensions and we shall call them non-extensional. As Lewis (1976) shows, most adjectives do not have extensions.

‡ So treated variables are of category e, and they have no lexical appearance in a language. The actual language variables, those introduced in section 2.1.1 are of category t/(t/e), and their intensions are rather sets of properties the things referred by the corresponding non-lexical objects in e have. We shall discuss the topic later.

introduced in section 2.1.1. and show its fully developed syntax and semantics.

2.1.3. Lambda-categorial languages

In this section we complete our description of formal categorial languages by introducing a concept of λ-categorial languages and defining a semantics for this extension. At the end of section 2.1.1 we discussed a specific example of an English sentence which escaped treatment in the pure categorial language framework. To justify the so-called strong sense derivation of the sentence

(1)   Every boy loves some girl

we wanted the expression

(2)   every boy loves x

to act as an operator for the phrase "some girl". We could not take (2) directly as an operator since we observed it belonged to the basic category t. But we can use the notion of λ-abstraction (see Appendix A) to say that

(3) $(\lambda x \,(\text{every boy loves } x))$

is a function that given an argument a yields the value

(3')   (every boy loves a)

Clearly. if (3) is to act as an operator in (1) it must fall into category IV $= t/e$. and therefore the bound-variable part of (3) belongs to a new derived category $(t/e)/t$. A rather obvious question comes to mind whether other elements of category IV can be represented in the form of (3). Indeed. it seems to be the case as $(\lambda x \,(\text{sleeps } x))$ and *sleeps* have the same linguistic significance in our grammar. Both expressions can be even regarded as just a redundant way to say the same thing i.e. *sleeps*: see also Cresswell (1973).

The introductory discussion above gives. I hope. enough intuition on how the λ-abstractions can be employed to extend the notion of pure categorial languages. We shall call the extension **λ-categorial**

languages.

In a brief introduction to λ-calculus (Appendix A) we have said nothing to prevent one from a more generalized use of λ-abstractions than those mentioned above. Indeed, as we shall see in a moment, the concept can be naturally used in other categories than IV. That is, whenever we have an expression α from a category τ, and a variable x from another category σ, then the λ-expression $(\lambda x \alpha) \in E(\tau/\sigma)$. If then x occurs free in α we can rewrite α as (βx) where $\beta \in E(\tau/\sigma)$. According to the suggestion above, we can associate with every member β of $B(\tau/\sigma)$ the λ-term $(\lambda x (\beta x))$. The two definitions below put the above considerations into formal rules. The definitions are derived from Cresswell (1973).

<u>Definition 5</u>

For any category $\sigma \in CAT$, let X_σ be an infinite set of variables (or linguistic pseudo-variables) of that category. Let further for any two $\sigma_1, \sigma_2 \in CAT$, $X_{\sigma_1} \cap X_{\sigma_2} = \emptyset$ if and only if $\sigma_1 \neq \sigma_2$. □

The next definition, although essentially from Cresswell (1973), has been rearranged to better suit our previous formal definition of pure categorial languages.

<u>Definition 6</u>

By a λ-categorial language L we shall mean an ordered 5-tuple $<BCAT, X, \Delta, B, E>$ such that

(i)   BCAT is a finite, nonempty set of basic syntactic categories.

(ii)   CAT is the set of all syntactic categories as described in section 2.1.1.

(iii) $\Delta = <\Delta_{\sigma_1}, ..., \Delta_{\sigma_k}, ...>$ is a sequence of finite, disjoint sets of which all but finitely many are empty. Δ will be called the lexicon of L.

(iv) $X = <X_{\sigma_1}, ..., X_{\sigma_k}, ...>$ is a sequence of infinite, disjoint sets of variables as in Definition 5.

(v)   B is a function from CAT such that for any $\sigma \in CAT$, $B(\sigma) = \Delta_\sigma \cup X_\sigma$.

(vi)   E is a function from CAT such that for any $\sigma, \tau, \sigma_1, ..., \sigma_n \in$ CAT

(a) $B(\sigma) \subseteq E(\sigma)$

(b)   if $\delta \in E(\tau/\sigma_1, ..., \sigma_n)$ and $\alpha_1, ..., \alpha_n \in E(\sigma_1), ..., E(\sigma_n)$ respectively, then $<\delta, \alpha_1, ..., \alpha_n>$
$\in E(\tau)$

(c)   if $\alpha \in E(\tau)$ and $x \in X_\sigma$ then $(\lambda x.\alpha) \in E(\tau/\sigma)$. $\square$

Definition 6 differs in one point from that of Cresswell (1973). I made sets X_σ to be parts of $B(\sigma)$ rather than $E(\sigma)$ which is consistent with our previous statements from section 2.1.1. As we can see, the definition of λ-categorial languages at their syntactic level is a merely straightforward extension of Definition 3 from section 2.1.1. Two new elements appear. First, the concept of a linguistic variable has been introduced, which may be regarded as a generalized concept of pronoun (or maybe better pro-σ for category σ) Second, the definition allows λ-terms of any sort to be among well-formed expressions of L. To be more specific, let us reconsider our grammar for a fragment of English which we introduced in section 2.1.1.

2.1.3.1. English as a Categorial Language - Revisited

Nothing will change in the base description of the grammar, except the fact that we can assume the existence of lexical variables among basic expressions of categories. Yet, the category e will remain empty, and no linguistic variables will be placed in B(e), precisely for the same reason we have voided this category from other lexical elements. Let us first consider a very simple example. The sentence *John runs* on Figure 2.3 has an obvious derivation in categorial grammar along the lines shown

No other valid derivation existed in the pure categorial version of the grammar presented in section 2.1.1. Now, however, in the light of (vi)(c) from Definition 6 we could get another derivation as well, as shown on Figure 2.4. This particular derivation poses no problems when regarded in general terms of λ-categorial languages. Moreover, it exhibits some extremely interesting properties of λ-derivation. But in our fragment of English the derivation is blocked by the fact that we have no lexical elements

t (John runs)

t/(t/e)   t/e
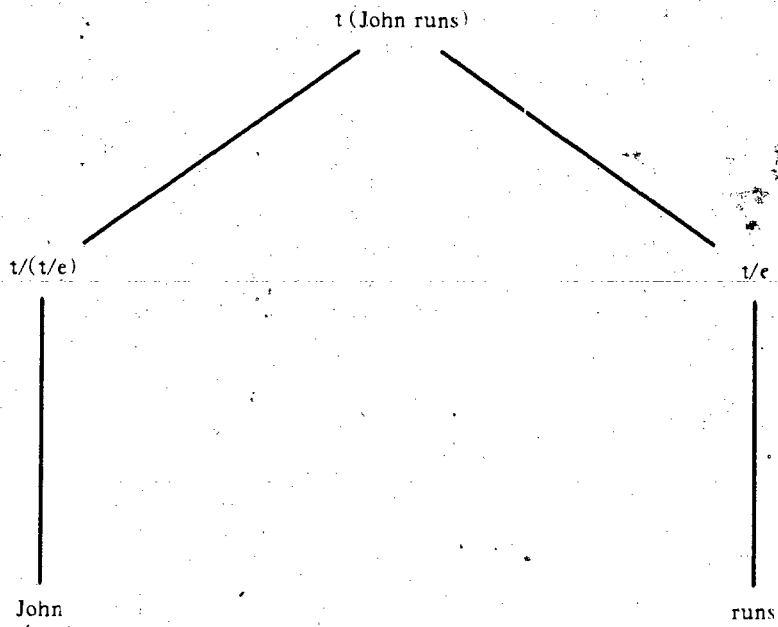
John   runs

Figure 2.3. A derivation of 'John runs' in the categorial grammar for English.
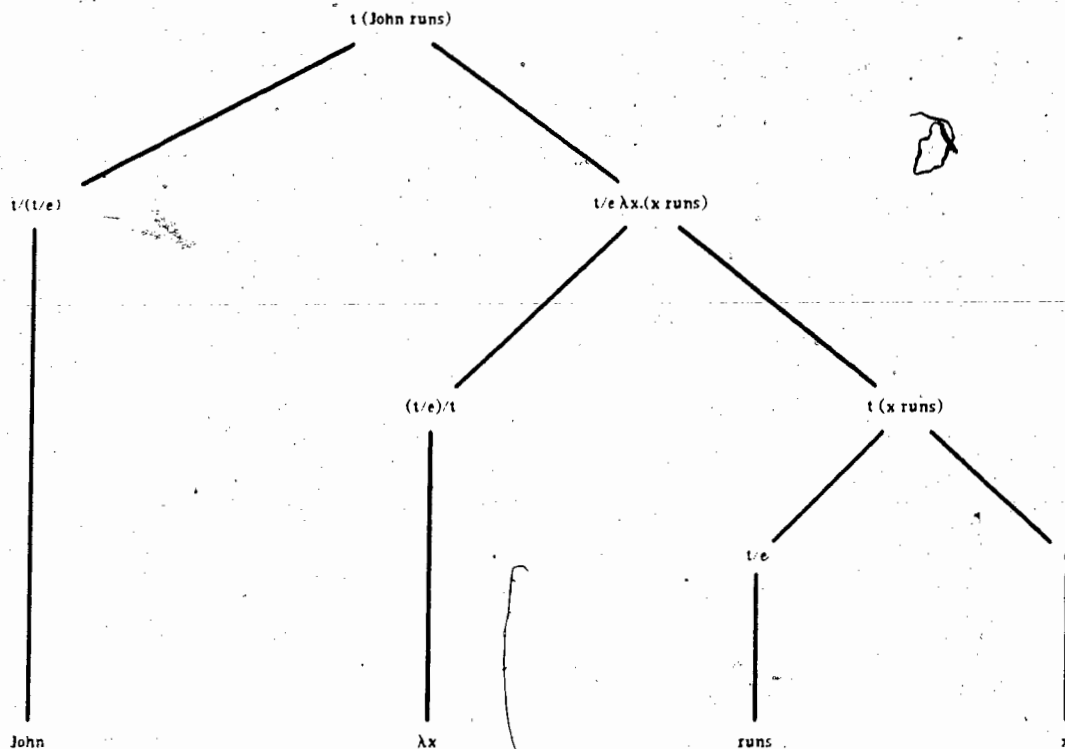
Figure 2.4. Another possible derivation.

within category e.

In the case of the sentence *Every boy loves some girl* the pure categorial approach gives an explanation for only the so-called *weak sense* (or *opaque)* reading of the sentence. which is well-known as ambiguous due to quantifier scoping. In our English fragment the weak sense derivation is depicted as Figure 2.5 (it is unique because of constraints stated above). Another valid derivation is still possible. however. yielding the so-called *strong sense* (or *transparent)* reading. Figure 2.6 illustrates this possibility. Observe that the derivation is also unique in this fragment.

We can imagine other circumstances when using λ-expressions is inevitable to explain some syntactic constructions (consider this sentence!). Montague (1974c-f). Thomason (1976). Partee (1976) are among those who offer examples. Because syntactic problems per se are not of interest to us. we
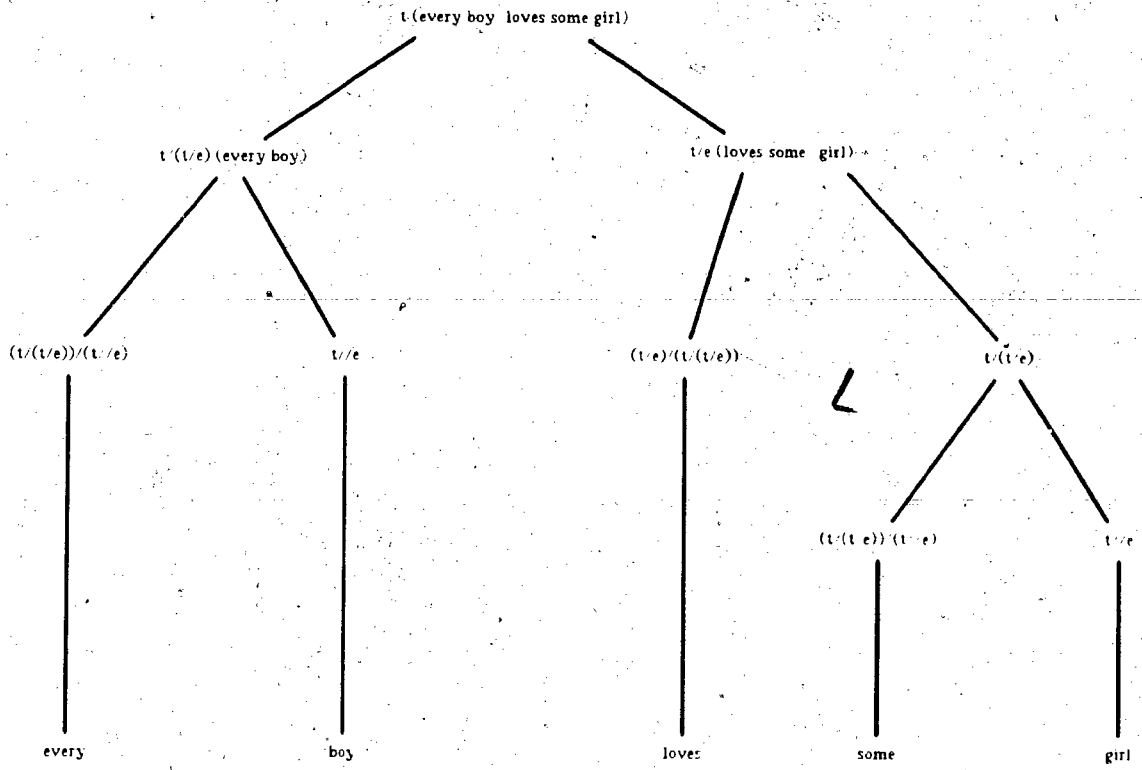
Figure 2.5. Weak-sense derivation for 'every boy loves some girl'
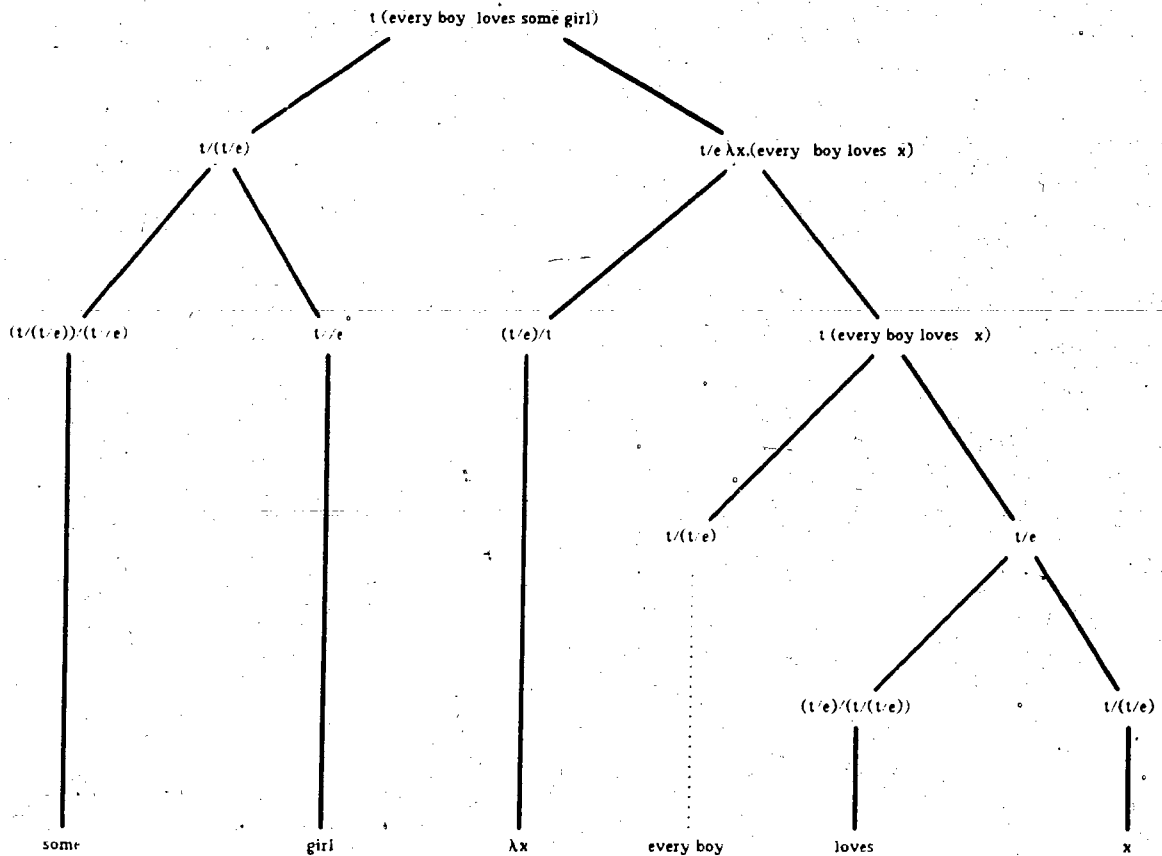
Figure 2.6. Strong-sense derivation of 'every boy loves some girl'.

shall not present any more sophisticated examples. and the interested reader is directed to the above works

We shall show. however. yet another example to see how the syntax of the λ-categorial grammar of our English fragment can explain the so-called referential reading vs. nonreferential reading (or *de re / de dicto*) ambiguity of some sentences. Let the sentence in question be *John wants to marry a queen*. Certainly. *a queen* ∈ B(T) = B(t/(t/e)). and *wants to* ∈ B(IV//IV) = B((t/e)//(t/e)). In one reading. John just wants any queen to become his wife. and therefore we can say of the **non-referential reading**. It will have the derivation tree as shown on Figure 2.7 below. On the other hand.

t (John wants to marry a queen)

t/(t e)

t/e (wants to marry a queen)

IV 'IV

t/e (marry a queen)

(t/e)/(t/(t/e))

t (t e)

John

wants to
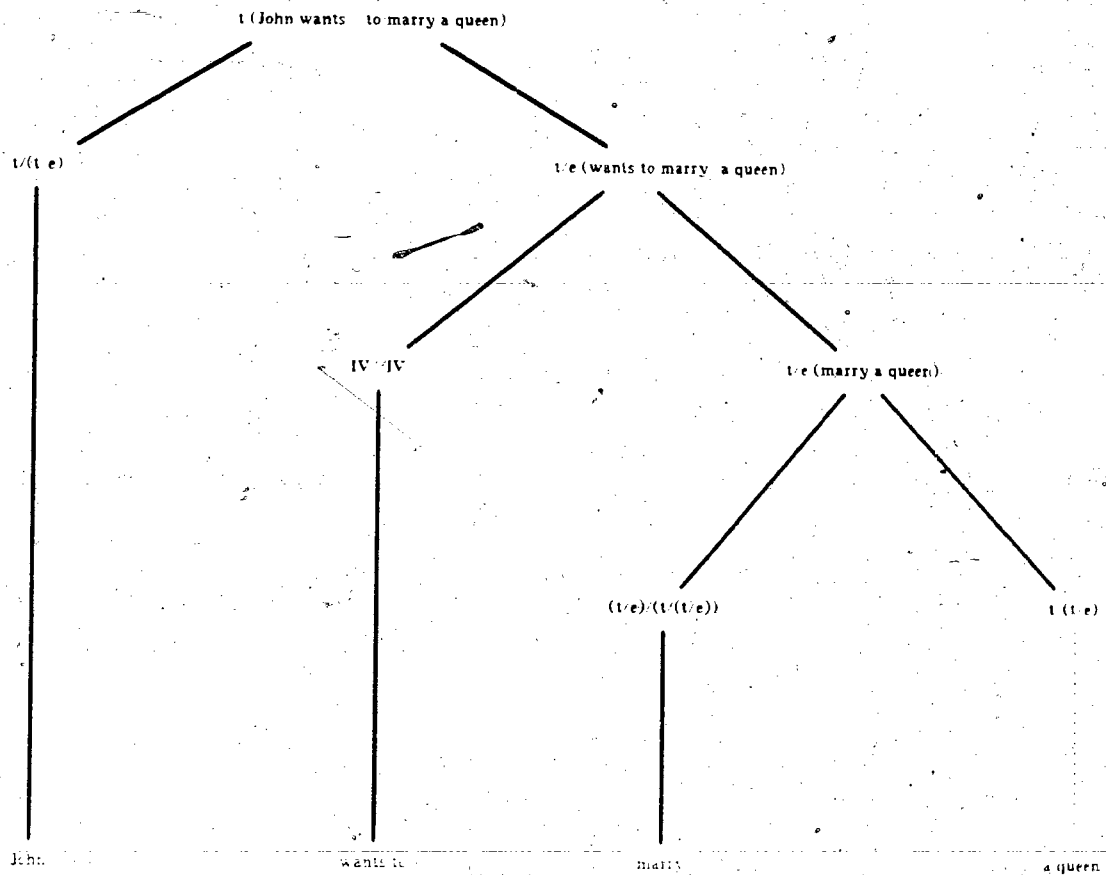
marry

a queen

Figure 2.7. Non-referential derivation for John wants to marry a queen

one could observe that John has a particular queen in mind, thus forcing the derivation depicted as Figure 2 8. These few examples should be enough to convince us of the remarkable expressive power of λ-terms among expressions of a categorial language. What remains to be seen is the semantic and pragmatic side of the λ-categorial languages. We shall turn to this issue now.

2.1.4. Semantics and Pragmatics of λ-categorial Languages

In this section we shall extend the notions of extension and intension over indices for λ-categorial languages, along the lines of the section 2.1.2. Then we shall turn to the even more vital question

t (John wants to marry a queen)

t/(t/e)

t/e λx.(John wants to marry x)

(t/e)/t

t (John wants to marry x)

a queen

λx

(essentially same as above)
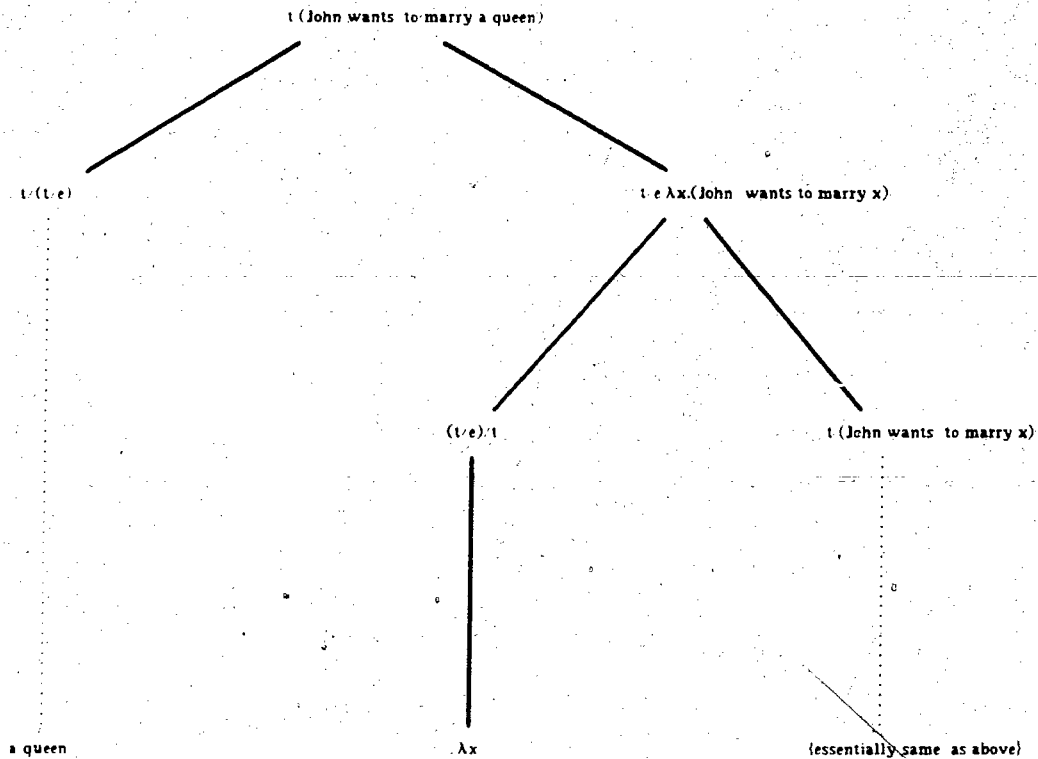
Figure 2.8. Referential reading derivation of John wants to marry a queen

when dealing with natural language meaning - pragmatics Ultimately we shall briefly discuss what we feel are the most important aspects of the meaning representation theory espoused by Montague (1974d-f) and still under development by his followers Thomason (1976), Partee (1976), Dowty (1976)

2.1.4.1. Semantics

For the clarity of presentation, let us observe first how the Cresswell's V operation (value assignment, meaning) has to be modified to cover two new types of expressions: linguistic variables and λ-terms. We do not need any changes to be done to domains D_σ of categories. The universe remains

the same. only the notation gets richer.

As "ordinary" elements of syntactic categories. linguistic variables are supposed to denote some "things" in respective domains. but unlike the former they cannot be regarded as referring to any particular generalized "things" in domains unless given in a specific pragmatic context. Following Cresswell (1973). we introduce a new function $v : X^+ \to D^+$ such that whenever $x \in X_\sigma$. $\sigma \in CAT$. then $v(x) \in D(\sigma)$. We shall call the function v **the variable binding coordinate** of value assignment V. Let us further define a function v_a^x which is much like v with only one exception. That, is for any $y \in X^+$. $y \neq x$. $v_a^x(y) = v(y)$. while $v_a^x(x) = a$. Finally we define our extended function V as V_v and say. after Cresswell (1973). that

(1)   for any symbol δ of L. δ is not a variable. we have $V_v(\delta) = V(\delta)$.

(2)   if $x \in X^+$ then $V_v(x) = v(x)$.

In this way we have just accommodated the variable binding coordinate into our earlier notion of meaning.

Definition 7 [based on Cresswell (1973)]

Let L be a λ-categorial language. and A be the set of all variable binding coordinates for X^+. The value assignment function V_v for L is defined as follows:

(i)   for any constant symbol α of L. and for any $v \in A$. let $V_v^+(\alpha) = V_v(\alpha)$.

(ii)   if $\alpha = <\delta. \alpha_1. \ldots \alpha_n>$. $\delta \in E(\tau/\sigma_1 \ldots \sigma_n)$ and $\alpha_1 \ldots \alpha_n \in E(\sigma_1) \ldots E(\sigma_n)$ respectively then for any $v \in A$

$$V_v^+(\alpha) = V_v^+(\delta)(V_v^+(\alpha_1). \ldots V_v^+(\alpha_n))$$

(iii)   if $\alpha = \lambda x.\beta$. $x \in X_\sigma$. $\beta \in E(\tau)$. then $V_v^+(\alpha) : D_\sigma \to D_\tau$. and for any $a \in D_\sigma$. $V_v^+(\alpha) = V_{v_a^x}^+(\beta)$ □

The above definition looks a little complicated. When compared to what we have said in section 2.1.2 the only new element appears to be (iii). Yet. this distinction quickly disappears when reconsidered in the context of λ-convertibility introduced in the last section. What is more important. however. still

seems to be hidden from the reader. To state it explicitly let us observe that introduction of variable binding coordinate allows us to determine meanings of so-called **open expressions** i.e. these in which at least one variable occurs free. A value of function V_v^+ can be no longer considered as a generalized "thing" from an appropriate domain unless

(a)   a variable binding coordinate v is chosen from A. and

(b) v is determined for every free variable in a given language expression which meaning is to be computed

To be even more specific. let us reconsider the domain D_0 of propositions which we chose to regard as generalized truth values. The open sentence *Every boy loves x* cannot be said to express any proposition which may or may not be true in some possible world until the meaning. or intension. of the variable x is determined. Thus we are ultimately arriving at our previous notion of index and what Lewis (1976) called its assignment coordinate. What once seemed to be quite mysterious. now helps us to maintain our concept of intension over indices unchanged. as introduced in section 2.1.2. In the case of λ-categorial languages we can still talk about intensions as generalized "things" in domains. where the assignment coordinate provides intensions for every variable employed in language. One can rewrite Definition 7 as $V_v^+(\alpha) = Int(\alpha)$ for any expression α of language L. What still remains uncertain is how to deal with λ-terms. Recall. however. the principles of λ-convertibility from the previous section It should be clear that the following will hold:

(i) $Int(\lambda x\,(\alpha x)) = Int(\lambda y\,(\alpha y))$ providing y does not occur free in α.

(ii) $Int(\lambda x\,(\alpha x)) = Int(\alpha)$.

(iii) $Int((\lambda x\,(\alpha x))(\beta)) = Int((\alpha\beta))$ providing $x, \beta \in E(\sigma)$. $\sigma \in CAT$. $x \in X_\sigma$.

What we now have in hand is the complete semantics of categorial languages. Let us summarize briefly For any syntactic category $\sigma \in CAT$ there is a domain $D(\sigma)$ of **generalized things** which. relative to the index. may or may not realize "real" things in some possible worlds. When we speak of things. we mean entities of any sort. Generalized things do not exist in reality (except when we

establish semantics of some metalanguage as for example the language of this thesis †). or in any

possible "reality". They could be rather thought of as ideas When considering semantics of our frag-

ment of English introduced in section 2.1.1. propositions from the domain D(t) were generalized truth

values: domain D(e) consisted of generalized names which for a given index may denote "real" objects

Domains of derived categories were sets of generalized functions between generalized things of com-

ponent category domains

Our semantics has been therefore based on rather mysterious out-of-this-world entities which

the generalized things actually are. But as with every powerful tool, the idea may carry some danger

when applied too literally to the description of a natural language meaning representation As we have

said. generalized things are partial functions from indices to possible worlds It may be the case that a

sentence uttered in some possible world. say in our present real world contains elements which do not

have extensions in this world. To be more specific. let us consider the following sentence given by

Cresswell (1973)

(1)   The present king of France is bald.

Clearly. the definite description *the present king of France* of category T   t/(t/e) lacks extension

in our real world. What would therefore be the extension of (1) in our world? It will be neither truth

nor falsity as we cannot find any object which is currently a king of France to verify whether he is bald

or not But (1) is perfectly meaningful. The next paragraph tries to shed some light on the problems

which are widely considered as pragmatics of a language

2.1.4.2. Pragmatics

In section 2.2 we have said that we would consider indices as complexes consisting of all neces-

sary coordinates to provide a unique (possibly) context in which expressions of a language are uttered

We followed Lewis' (1976) proposal that they should include. among others. the possible world coordi-

---

† However we get another level of generalized things which cannot be objects of the language

nate. and a collection of contextual coordinates like the time coordinate. the place coordinate. the speaker coordinate. the audience coordinate. etc. It should be relatively clear that one can state no upper limit on the number of contextual coordinates necessary to validate an utterance in general. If this is the case. why not have coordinates for country. religion. climate. or even previous drinks as suggested by Cresswell (1973)? This approach seems to require some advance specification of all coordinates and that this knowledge would be crucial to understand language utterances in general. This assumption may be valid to some degree. as for example if John Smith says *I am hungry* it will be understood that in a possible world w where John Smith exists. and with the speaker coordinate s set to him. the utterance expresses the proposition that John Smith is hungry. The proposition will be true at an index i which contains w and s as its possible world and speaker coordinates respectively if John Smith is indeed hungry at i. and may be false at other contextual coordinates. On the other hand. if you find the sentence typed on a sheet of paper in a typewriter you will probably **presume** that somebody. whoever he was. expressed the proposition about himself. You may wonder whether the proposition is (was) true or not on the basis who the utterer could be. Yet. you **understand** the sentence. more or less. even if you lack any contextual knowledge on how the event of typing took place. If the line was printed by a monkey which hit random keys of the typewriter then your earlier presumption was wrong. and the sentence cannot be validated as true or false. It simply has no **sense**.

The same type of consideration underlies an analysis of the sentence cited in the last paragraph. You understand what it **means** even if you do not share the speaker presupposition that there is a unique king of France. If the speaker made his statement seriously. he probably believed that a king of France actually exists. and accordingly. he believed he expressed a proposition. For us. however. the sentence expresses no proposition. and one may answer: "This makes no **sense**. there is no present king of France."

We may continue the discussion to great length. but it should be becoming clear now that what we have called semantics of a language actually consists of two distinct levels. The level which operates on generalized objects and treats **meanings** which we shall call **semantics**. and the level

which validates generalized objects in some context and which we shall call **pragmatics**. see Montague (1974e). Stalnaker (1972). Cresswell (1973). The key observation to make is that the contextual coordinates of index are actually independent of meanings. If supplied they may perhaps establish **sense** (if there is one) of a language expression. Let us divide our indices into two distinct sets, say C and N. of complexes which coordinates are. and are not context dependent. respectively. Now, let α be any expression of type 0 of a λ-categorial language L. When considered only with respect to N (let us call it the set of possible worlds. but remember that N can consist of multi-coordinate tuples, not only of our previous possible world coordinate) some elements of α may, in general, remain unspecified in the absence of contextual information. These elements are generally free variables in α. In natural language they include pronouns (free occurrences). demonstrative pronouns. relative time specifications such as "today". "tomorrow". etc. The interpretation of α with respect to N and the set D^+ of all domains will not be a proposition because its truth value cannot be determined until all free variables of α are given values. But this will be a function which when supplied with proper contextual (or pragmatic) information will yield a proposition. Cresswell (1973) calls these species **open propositions**. Open propositions are therefore functions from C into D_0 cf Cresswell (1973). Stalnaker (1972)   .

This notion is still very informal but can be easily extended on categories other than 0 (or t). We shall formalize it now. along the lines of Montague (1974e).

Definition 8

Let σ be any one category of a λ-categorial language L. Let N be a set of all possible worlds. and C a set of all contexts of use. As earlier. let $D(\sigma)$ be a domain of all generalized things of type σ. Let further $\Delta(\sigma)$ be a set of all possible denotations of type σ. i.e. the set of all non-empty extensions of generalized things from $D(\sigma)$ in all possible worlds. We shall call the set $S(\sigma) = \Delta(\sigma)^N$ of all functions from possible worlds into the set of possible denotations of type σ. the set of **senses** of expressions of category σ. Accordingly we shall call the set $M(\sigma) = \Delta(\sigma)^{N \times C}$ of all functions from indices into the set of possible denotations of type σ. the set of **meanings** of expressions of category σ. $\square$

One can easily observe that with respect to our previous discussion $\Delta(\sigma)^{N \times C} = D(\sigma)$ i.e meanings are simply generalized things (putting aside for a moment Lewis' structural meanings). This is consistent with our earlier statements that meanings are intensions (or complexes of intensions, according to Lewis). By distinguishing semantic and pragmatic contexts N and C (which previously created indices $I = N \times C$) we have obtained a new level of language - the pragmatic level of senses. We can say therefore that while semantics deals with meanings of language expressions, pragmatics concerns their senses. Both of these levels of language, though distinct, remain in a close relationship. Observe that

$$M(\sigma) = \Delta(\sigma)^{N \times C} = (\Delta(\sigma)^N)^C = S(\sigma)^C.$$ Meanings are therefore functions from pragmatic contexts into senses. For category t of sentences $M(t) = S(t)^C$ is then the aforementioned set of open propositions. In general, one can say meanings of expressions of category σ are **open generalized things** of type σ. In this approach, to get the meaning of a language expression of category α we first construe the set $s(\alpha)$, $s \in S(\sigma)$, of all possible senses for that expression and then validate it in given context. Another way would be to establish the context first i.e to make an open sentence become closed under some interpretation, and then look at its extensions in all possible worlds. cf. Stalnaker (1972), i.e

$$M(\sigma) = (\Delta(\sigma)^C)''$$

Returning to the example sentence

(1)   *The present king of France is bald*

we can say that it has a meaning. It is that open proposition (i.e function from C to $S(t)$) which in a given context may or may not become a proposition, i.e an element of $D(t)$. It has, however, no sense in any possible world $w \in N$ in which the phrase *the present king of France* lacks a denotation in $\Delta_w(T)$, which is the case in our "real" world. †

The above discussion does not exhaust all problems concerned with meaning analysis of λ-categorial languages in general, and natural languages in particular. But it gives a good account of the complexity of these problems. Recognizing the distinct pragmatic level complicates the analysis of

---

† No generalized thing satisfying the given definite description has a denotation in $\Delta_w(T)$.

natural language so significantly that some philosophers and logicians. see for instance Montague (1974d) and Lewis (1976). often disregard this distinction by putting both semantics and pragmatics under the common banner of semantics. or indexical semantics. in our previous sense. This action would oversimplify some problems as Cresswell (1973) and Stalnaker (1972) point out, yet it makes difficult problems more amenable to analysis.

No more general theory of languages which could be relevant in the context of this work pays so much attention to the problems regarding pragmatics. The pragmatic side of language analysis is mostly considered as a future battlefield. This is the case in such general linguistic theories as Standard Theory. Chomsky (1965): Generalized Phrase Structure Grammars. Gazdar (1979): Lexical Functional Grammars. Kay (1979): or Definite Clause Grammars. Colmerauer (1978). Warren (1980): and variations of them. These facts suggest strongly that the approach presented herein appears to be a promising step toward a universal theory of language and meaning.

2.1.5. Attempts to design a meaning representation

Thus far in this chapter we have introduced a formal definition of categorial languages and then showed how to design semantics and pragmatics for them. Based on the working assumption that natural languages are just special cases of categorial languages we saw. through selected examples. how the notions of semantics and pragmatics worked on English. We introduced the semantics directly. i.e we showed how a language expression can be assigned meaning and/or sense. Yet; we do not know how to represent the meanings of language expressions. The notions of intension and extension. though very helpful and intuitively acceptable. still escape any tractable empirical (hence computational) treatment. What we need is some formal notational system that would allow for representing and manipulating these and other aspects of meaning. We propose an intermediate level between a language (its syntax and sentential form) and meanings. which we shall call a **meaning representation**. If we can design a meaning representation level as another linguistic system. it will be assumed that the relationship between well-formed expressions of this meaning representation language and the

meanings they represent is significantly simpler than the corresponding relationship between the original language and the realm of meanings. The main goal of this research is to design a proper meaning representation for natural languages, and this will be discussed in following chapters. There are at least several possible ways to think of the meaning representation level, and we shall postpone this consideration to Chapter 3. Now, we would like to present briefly one specific approach to treat semantics of a fragment of English where the meaning representation level has been introduced in the form of the formal logical apparatus of intensional logic. We are referring to Montague's (1974f) PTQ, also described by Dowty (1981).

Any indirect introduction of the semantics of a language with a meaning representation level should consist of three main steps (this is indeed the case in PTQ):

(i)   choosing some formal, simple, artificial language to be the **meaning representation language**;

(ii)   designing the semantics of that language, and

(iii)   working out a set of rigorous rules of translation from source language into the meaning representation language.

The meaning representation language may contain more or fewer semantic elements in it, but certainly it must be rich enough to preserve meanings (and senses) of source language expressions. It would be advantageous, to propose even several intermediate meaning representation levels with those intermediate levels closer to the source language containing less semantic elements. We shall discuss this topic in some length in Chapter 3.

In PTQ, Montague chose intensional logic as the meaning representation language for his fragment of English. Intensional logic has a very semantic character and such a meaning representation language lies more on side of meanings than the source language. It contains, in addition to ordinary logical symbols and operators, temporal and modal operands, and what makes it more akin to meanings themselves: the notions of intension and extension built into it. The following presentation will be probably incomplete, but our goal in this paragraph is to emphasize the method without going into the

detail of this particular solution for some particular fragment of English.

Consider again the fragment of English introduced in section 2.1.1, which is based on the PTQ's fragment. There are some minor syntax differences for that fragment compared to that of PTQ. Montague limited his presentation to only those syntactic categories we mentioned explicitly when defining sets of basic expressions for them. These are: t, e, t/e = IV, t/(t/e) = T, IV/T = TV, IV/IV = IAV, t//e = CN, t/t, IAV/T, IV/t, and IV//IV. Refer to section 2.1.1 for their detailed description. For any other category A different than the above, Montague assumed that it has no basic elements, i.e. $B(A) = \varnothing$. What follows from this assumption is, among others, that the category $(t/(t/e))/(t//e)$ of quantifiers has no basic expressions, and Montague treats quantifiers like *every, the, a (an)*, etc, as distinguished operators. Their role in grammar, as well as the role of other non-classified expressions (as *such that*) is explained in a number of syntactic rules accompanying the grammar. The rules are also responsible for formulating morphologically and structurally well-formed expressions in the fragment with the proper distribution of tense, person, number and gender information, as well as pronominal references. These particular details should not concern us as long as we discuss the method in general. It will also be unimportant that Montague used his grammar in a generative rather than an analytic manner. Also since the details of morphologically well-formed expressions do not bother us, we shall normally neglect these syntactic rules of PTQ which are concerned with this problem. It is important to remember that this fragment can be regarded as a λ-categorial language with all consequences of this assumption. It should be more perspicuous if we think of expressions of the fragment as of their phrase-markers.

We assume that the syntax of the fragment of English is properly understood. Now we shall turn to the description of PTQ's meaning representation language.

In PTQ, the meaning representation language is the tensed intensional logic (IL). The appropriate definitions are given below, along the lines of Montague (1974f).

<u>Definition 9</u>

Let t, e, s be three distinct, fixed, atomic symbols. By the set of **types** it will be understood the smallest set TYPES such that:

(i) e, $t \in$ TYPES.

(ii)   for any two a, $b \in$ TYPES, $<a.b> \in$ TYPES.

(iii) for any $a \in$ TYPES, $<s.a> \in$ TYPES. $\square$

We may informally regard t and e as basic types, any pair $<a.b>$ as a compound (or derived) type, and s as a "type" of indices. Observe, s is not in TYPES.

The coincidence of the symbols used for basic types in IL and those denoting basic categories in the categorial grammar for the fragment is not accidental. Indeed, as we shall shortly see, they stand in one-to-one correspondence. Thus in IL, e is the type of individuals, while t is the type of truth values. For compound types, the triangle bracket notation parallels the slash notation in categorial grammar. In fact, both symbolisms are strictly equivalent, and the distinction made by Montague had been intended for exposition purposes only. What appears a little unfortunate is that the order of constituents gets reversed. Thus a type $<a. b>$ denotes the objects which when combined with an object of type a, give an object of type b. In other words, objects of type $<a. b>$ are functions from objects of type a to objects of type b. The definition states that for any type a, there is the *intensional* or *generalized* type $<s. a>$ that classifies functions from indices (s) into objects of type a. Observe that although the type $<s. a>$ consists of generalized objects of type a, they are treated as atomic concepts in IL. This is because no expression of this type can ever be a functor in IL, as s is not a type itself. This move made intensionality transparent in IL, just as it appears in natural language.

Thus $<s. e>$ id the type of individual concepts, or generalized individuals in our terminology. If an individual j is of type e then its intension $\hat{j}$ is of type $<s. e>$. Similarly, the type $<s. t>$ is that of propositions, or generalized truth values. Other types are composed in the usual way. Some of

them, however, are of particular importance to our discussion. If a is a type of IL then $<a, t>$ is a type of sets of objects of type a. This notion closely corresponds to the extensional one-place relations on objects of type a. Subsequently, the type $<e, t>$ denotes extensional one-place relations on individuals, and $<<s, e>, t>$ classifies such relations on individual concepts. This gives us a straightforward method to generate sets of sets, sets of sets of sets, etc. For example, the type $<<e, t>, t>$ is the one of sets of sets of individuals, that is, the type of extensional one-place relations on sets of individuals.

Intensions of sets are called *relations-in-intensions*. The one-place relations-in-intension are called properties. If a is any type of IL then the properties of objects of this type are classified into the type $<s, <a, t>>$. For example the type $<s, <e, t>>$ contains properties of individuals, while the type $<s, <<s, t>, t>>$ denotes properties of propositions. A step up in the hierarchy gives us properties of properties, and so on. The type $<s, <<s, <e, t>>, t>>$ is the one of properties of properties of individuals.

Many-place relations cannot be identified with sets but the representation of appropriate types is equally straightforward. Thus, for any two types a and b in IL, the type $<a, <b, t>>$ is the one of extensional two-place relations between objects of type b and objects of type a. In the same way we can define types of 3, 4, and more-place relations. For instance, the type $<e, <e, t>>$ classifies extensional two-place relations on individuals. Then the type $<s, <e, <e, t>>>$ denotes two-place relations-in-intension on individuals. In general the two-place relations-in-intension between objects of types b and a fall into the type $<s, <a, <b, t>>>$. Both a and b can be quite elaborate. For instance, the type $<s, <<s, t>, <e, t>>>$ is the one of two-place relations-in-intensions between individuals and propositions.

Not all types of IL are of equal importance for classifying expressions of natural language. As we shall shortly see, the proposed correspondence between categories of the fragment and the types of IL leaves many otherwise meaningful types virtually untouched. Before we turn to a discussion of the

problems of translation, however, we have to consider the syntactic side of IL.

Definition 10

For any type $a \in$ TYPES there exist two infinite sets Var_a and Con_a of variables and constants of type a. By $v_a^n \in Var_a$ we shall mean the nth variable of type a. $\square$

Beside variables and constants each type will consist of well-formed (or meaningful) expressions.

Definition 11

For any type $a \in$ TYPES let WFE_a be a set of well-formed expressions For any type $a \in$ TYPES let WFE_a be a set of well-formed expressions

(i) $Var_a \subseteq WFE_a$

(ii) $Con_a \subseteq WFE_a$

(iii)   if $\alpha \in WFE_a$, and $x \in Var_b$, then $\lambda x.\alpha \in WFE_{<b.a>}$ †

(iv)   if $\alpha \in WFE_{<a.b>}$, and $\beta \in WFE_a$, then $\alpha(\beta) \in WFE_b$

(v)   if $\alpha, \beta \in WFE_a$, then $(\alpha = \beta) \in WFE_t$

(vi)   if $\phi, \psi \in WFE_t$, $x \in Var_a$ for any one $a \in$ TYPES, then also the following are members of
WFE_t: $\neg\phi$, $(\phi \& \psi)$, $(\phi \cup \psi)$, $(\phi \rightarrow \psi)$, $(\phi \equiv \psi)$, $\forall x\phi$, $\exists x\phi$, $\square\phi$, $\uparrow\phi$, $\downarrow\phi$

(vii)   if $\alpha \in WFE_a$ then $(\hat{\ }\alpha) \in WFE_{<s.a>}$

(viii)   if $\alpha \in WFE_{<s.a>}$ then $^v\alpha \in WFE_a$ $\square$

We shall consider the sum $\bigcup\limits_{a \in TYPES} WFE_a$ as the set of all well-formed (or meaningful) expressions of intensional logic. Some comments are necessary in order to explain the meaning of symbols used in definition above. For a variable x of type a, the expression $\lambda x.\alpha$ is understood as denoting that function from objects of type a which takes as value (βx) if $\alpha = \beta(x)$, or α if x does not occur free in α. Any expression $\alpha(\beta)$ denotes the value of the function denoted as α on the argument denoted by β.

---

† Observe the opposite ordering of derived type tuples in comparison to categories of a categorial grammar.

The logical symbols: $=$, $\neg$, &, $\cup$, $\rightarrow$, $\equiv$, $\forall$, $\exists$ are understood to have their usual interpretations. $\Box$ may be read as *it is necessary that*; $\uparrow$ as *it will be the case that*; $\downarrow$ as *it has been the case that*. The expression $(^\wedge\alpha)$ denotes intension of α; $^\vee\alpha$ denotes extension of α. Observe that $^\vee\alpha$ is meaningful only if α denotes some generalized thing i.e intension or sense.

The system of intensional logic is prepared so that it can deal with two classes of things. Expressions of intensional logic can include references to generalized things (by using intensions), extensional things, and any mixture of them. As the actual object base for the formalism, Montague introduced the set A of entities (or individuals, or possible individuals) which we can identify with $\Delta(e)$ from before, with respect to the set $N = W \times T$ of possible worlds, where W determines the possible world coordinate, and T sets moments of time. Accordingly, he introduces the set of possible denotations for any type a, with respect to W and T. Thus we have:

(1) $\Delta(e) = A$

(2) $\Delta(t) = \{0, 1\}$

(3) $\Delta(<a.b>) = \Delta(b)^{\Delta(a)}$

(4) $\Delta(<s.a>) = \Delta(a)^{W \times T}$

Observe that $\Delta(<s.a>) = S(a)$ is the set of **senses** of expressions of type a, as introduced earlier.

Montague introduces semantics of intensional logic in two steps.

Definition 12

By an **interpretation** U of intensional logic we shall mean a 5-tuple $<A \ W. \ T. \leqslant. \ F>$ such that:

(i) A, W, T are non-empty sets.

(ii) $\leqslant$ is a linear ordering over T.

(iii) F is a function from $\bigcup_{a \in TYPES} Con_a$ and such that for any $a \in TYPES$, $\alpha \in Con_a$, we have

$F(\alpha) \in S(a)$. $\Box$

An interpretation sets, as we can observe, the pure semantics of the intensional logic. What we still need is some notion of pragmatics to account for meanings of arbitrary expressions. Let then g be a function from all variables i.e from $\bigcup_{a \in TYPES} Var_a$, such that for any $x \in Var_a$, $g(x) \in \Delta(a)$. Therefore g plays the role of assignment coordinate, and is understood as setting the pragmatic context for interpretation. Let us call such an interpretation U_g. Accordingly, let Int_{U_g} and $Ext_{U_g}^{w,t}$ be two functions whose domain is the set of all well-formed expressions of intensional logic. We shall call Int_{U_g} and $Ext_{U_g}^{w,t}$ intension and extension respectively. Whenever it is not ambiguous we shall drop the subscripts of the symbols Int and Ext for short. As usual, when $<w, t>$ is a point of reference then for any $\alpha \in WFE^+$, $Int(\alpha)(<w, t>) = Ext^{w,t}(\alpha)$. The following definition establishes a full semantics (indexical semantics) for intensional logic.

Definition 13

Let U, g be an interpretation, and assignment coordinate for intensional logic. The semantics of the model according to U and g is described as follows:

(i)   if $\alpha \in \bigcup_{a \in TYPES} Con_a$ then $Int(\alpha) = F(\alpha)$, $Ext^{w,t}(\alpha) = F(\alpha)(<w, t>)$

(ii) if $\alpha \in \bigcup_{a \in TYPES} Var_a$ then $Ext^{w,t}(\alpha) = g(\alpha)$

(iii) if $\alpha \in WFE_a$, $x \in Var_b$ then $Ext_{U_g}^{w,t}(\lambda x.\alpha)$ is that function h from $\Delta(b)$ such that if $\beta \in \Delta(b)$

then $h(\beta) = Ext_{U_{g'}}^{w,t}(\alpha)$ where g' is as g with one possible difference that $g'(x) = \beta$

(iv) if $\alpha \in WFE_{<a,b>}$ and $\beta \in WFE_a$ then $Ext^{w,t}(\alpha(\beta)) = Ext^{w,t}(\alpha)(Ext^{w,t}(\beta))$

(v) if $\alpha, \beta \in WFE_a$ then $Ext^{w,t}(\alpha=\beta)$ is 1 (true) if and only if $Ext^{w,t}(\alpha) = Ext^{w,t}(\beta)$

(vi) if $\phi \in WFE_t$ then $Ext^{w,t}(\neg\phi)$ is 1 if and only if $Ext^{w,t}(\phi) = 0$

Similarly, for other ordinary logical symbols: $\&$, $\cup$, $\rightarrow$, $\equiv$.

(vii) if $\phi \in WFE_t$, $x \in Var_a$ then $Ext_{U_g}^{w,t}(\exists x\phi) = 1$ iff there exists $\beta \in \Delta(a)$ such that

$Ext_{U_{g'}}^{w,t}(\phi) = 1$ where g' is as in (iii).

Similarly, for $\forall$ quantifier.

(viii) if $\phi \in WFE_t$ then

$Ext^{w,t}(\square\phi) = 1$ iff for any $w' \in W$, $t' \in T$, $Ext^{w',t'}(\phi) = 1$.

$Ext^{w,t}(\uparrow\phi) = 1$ iff for some $t' \in T$ such that $t \leqslant t'$, $t \neq t'$, $Ext^{w,t'}(\phi) = 1$.

$Ext^{w,t}(\downarrow\phi) = 1$ iff for some $t' \in T$ such that $t' \leqslant t$, $t' \neq t$, $Ext^{w,t'}(\phi) = 1$

(ix) if $\alpha \in WFE_a$ then $Ext^{w,t}(\hat{}\alpha) = Int(\alpha)$

(x) if $\alpha \in WFE_{<s,a>}$ then $Ext^{w,t}(\check{}\alpha) = Ext^{w,t}(\alpha)(<w,t>)$. $\square$

Some comments to avoid any confusion in this point. One should not mix the notions Int and Ext with logical symbols $\hat{}$ and $\check{}$ respectively. The latter are used to represent the meaning of source language. The former are parts of the semantics of the meaning representation language itself

The notions of truth and satisfaction for formulas of intensional logic are introduced below

Definition 14

If $\phi \in WFE_t$ then

(i) ϕ is satisfiable with respect to U, w, t if for some assignment coordinate g $Ext^{w,t}_{U_g}(\phi) = 1$

(ii) ϕ is true with respect to U, w, t if for any assignment coordinate g $Ext^{w,t}_{U_g}(\phi) = 1$ $\square$

In light of the above definitions elements of $WFE_{<a,t>}$ will denote 1-place relations on objects of type a. while elements of $WFE_{<a<b,t>>}$ will denote 2-place relations between objects of type a and objects of type b. Therefore elements of $WFE_{<s,<a,t>>}$ will denote **generalized** 1-place relations on objects of type a. i.e **properties** of these objects. cf. Montague (1974f). Whenever $\gamma \in WFE_{<s,<a,t>>}$. and $\alpha \in WFE_a$. then $(\check{}\gamma)(\alpha)$ asserts that α has the property of γ (at some point of reference $<w,t>$). Because γ is of type $<s, <a, t>>$. $\check{}\gamma$ is of type $<a, t>$. i.e it denotes a set of objects of type a. Thus to say that α of type a has the property γ is equivalent to the fact that α belongs to the set $\check{}\gamma$ at some index value $<w, t>$.

No more details of PTQ's intensional logic should concern us in this point as it would involve the full presentation of Montague's concepts. The next step is to establish some translation system to map source language expression (here. the fragment of English) into intensional logic. The general rules are simple and elegant.

Definition 15

Let f be a function from CAT into TYPES. We shall call f a mapping from a categorial language L into intensional logic if

(i) $f(e) = e$

(ii) $f(t) = t$

(iii) for any A. B ∈ CAT $f(A/B) = f(A//B) = <<s. f(B)>. f(A)>.$ □

The idea of the translation system such stated is to establish direct one-to-one correspondence between categories of a source language and the types of intensional logic. along the lines that for any category A expressions of this category translate into expressions of type $f(A)$. To account for translation of derived categories as in (iii) we must take into account the specifics of intensional logic semantics as designed for PTQ. Observe. therefore. that any expression of a derived category A/B is interpreted as denoting a function from intensions of expressions of category B (as they translate into $f(B)$) into expressions of category A (again. to their translations in intensional logic). Let then α be any expression of category A/B. and β be from $E(B)$. Therefore. α translates into $\alpha' \in <<s.$ $f(B)> f(A)>.$ while β translates into some $\beta' \in f(B)$. Applying α to β in categorial grammar we have got $\alpha(\beta) \in E(A)$. The intensional logic representation of this process will be the expression $\phi = \alpha'(\beta') \in WFE_{f(A)}$

To explain this translation we have to digress a little. It has been observed for some time that the linguistic objects appearing at functor-positions (that of category A/B) fall roughly into two. not disjoint classes. One. called intensional. requires that at the semantic level their elements take as

arguments intensions rather than extensions of appropriate type. The other. called extensional. contains all ordinary extensional operators. The observation was made in connection with such intensional verbs like *seek. believe. rise* or *change*. To guarantee the proper representation of meaning of such sentences like α *believes that* or α *rises* we have to translate them as *believe* ˊ(α. ˆφ) and *rise* ˊ(ˆα) respectively (observe that *believe* is intensional with respect to object position only). The treatment is assured in PTQ by converting the categories *IV/t* and *t/e* containing *believe that* and *rise* into the types < <s. t>. f(IV)> and < <s. e>. t> respectively. Similarly. the category *t/t* containing the intensional modal operator *necessarily* has been transformed into the type < <s. t>. t> so that the expression *necessarily* φ gets translated into *necessarily* ˊ(ˆφ).

One of the ingenious moves of Montague Grammar was the unification of the intensional and extensional classes by accommodation of the latter into the former. The superiority of the intensional class was additionally supported by the apparent examples of intensional intransitive verbs like *rise* or *change*. or such nouns as *price* and *temperature*. the category that has long been considered as purely extensional. The move allowed. inter alia. for the elegant solution to the so-called *temperature problem* refer. for instance. to Montague (1974f). Partee (1972). and Dowty (1981). The intensionality assumption created. as one may expect. some complications to the treatment of the truly extensional constructions. We still needed "ordinary" intransitive verbs and nouns like *walk* or *fish* to denote one-place relations. or sets. rather than properties. We also needed transitive verbs to be extensional on subject position. To ensure this. Montague restricted the intensional logic of PTQ by a number of so-called *meaning postulates*. They imposed some constraints on interpretation of some IL expressions providing them with extensional significance. The two postulates we are referring to here are numbered 2 and 3 in Montague (1974f). and they state that the following formulas are logically true in the PTQ's IL system.

Postulate 2

$\forall x \Box [\delta(x) \to \exists u \; x = \hat{}u]$ where δ translates any member of $B_{t//e}$ other than *price* or *temperature*.

Postulate 3

$\exists M \forall x \Box [\delta(x) \equiv {}^{\vee}M({}^{\vee}x)]$ where δ translates any member of $B_{t/e}$ other than *rise* or *change*.

In Postulate 2 variable x is of type $<s, e>$. while u is of type e. In the unconstrained translation δ falls into the type $<<s, e>, t>$. that is. the set of functions from indices to individuals. while only sets of individuals are needed. The postulate ensures the functions in translations of "ordinary" common nouns be constant functions. i.e have the same extensions at all indices. so that they can be identified with these extensions. i.e with sets of individuals. The second postulate. where M is of type $<s, <e, t>>$ and x is of type $<s, e>$. says that the truth conditions of sentences involving extensional intransitive verbs are equivalent to the membership of an individual (${}^{\vee}x$) in a set of individuals (${}^{\vee}M$). as required. To provide a convenient notation for extensional translations. Montague introduced the $\delta \cdot$ symbol to the effect that if $\delta \in WFE_{f(IV)}$ then $\delta \cdot$ assumes form $\lambda u \delta(^{\wedge}u)$. where $u \in Var_e$. The meaning of $\delta \cdot$ has been determined by the logical truth of the following formula:

- $\Box[\delta(x) \equiv \delta \cdot ({}^{\vee}x)]$. if δ translates any member of B_{CN} or B_{IV} other than *price. temperature. rise. or change*

The similar move was taken towards extensional transitive verbs.† The question naturally arises why we have to go into all these troubles with intensional translation of categories t/e and $t//e$ just for a handful of "extraordinary" elements. Then by adding special meaning postulates we exclude the majority of "ordinary" elements. rather than make just the opposite decision. Indeed. an alternative translation system has been suggested by Bennett (1976). His category-type equivalence scheme. where the categories t/e and $t//e$ translate into the type $<e, t>$. simplified. many aspects of IL representation. Unfortunately. the system proved weaker than original PTQ as. for example. the "temperature" puzzle cannot be properly solved there. See also Dowty et al. (1981) for discussion and comparisons.

---

† For postulates concerning transitive verbs. and numbered 4 to 7 the reader is referred to the original work. The postulate 1 ensures that all names used in the fragment have rigid designations.

With the mapping from categories into types so described. Montague sets a number of specific translation rules for expressions of the fragment of English into expressions of intensional logic. The rules parallel in some sense the syntactic rules assisting the grammar. The rules define a full transla- tion system for the fragment so that the main principle of meaning representation level (i.e its preserv- ing of semantics of source language) is maintained. These rules state, among others, that except category T of pseudo-names all basic expressions but *be* and *necessarily* translate into constants in intensional logic. In other words, for any $\alpha \in B_A$, $A \in CAT$, $A \neq T$, $\alpha \neq be$, *necessarily* α translates into a constant in $Con_{f(A)}$. *Be* and *necessarily* are regarded as second order operators and get a spe- cial treatment in PTQ.

Let us stop for a while to look closer at the semantics of *be*, mainly because it touches a very sensitive problem of logical equivalence. The translation of the verb *be* into PTQ IL has been designed to uniformly cover both the "identity" *be* as in *Morning Star is Evening Star*, and "predication" *be* as in *John is a man*. The translation is quoted from Montague (1974f)

- $be \rightarrow \lambda P \lambda x \ 'r'(\hat{} \lambda y(\ 'x = \ 'y))$

The formula says that the meaning of the "identity" *be* guaranties only that the two individual con- cepts meet at some common value of index, i.e they have common extensions. Here $P \in Var_{\langle s, \langle \langle s, \langle e, t \rangle \rangle, t \rangle \rangle}$, $x, y \in Var_{\langle s, e \rangle}$. Such an interpretation of identity *be* has important consequences for applicability of the substitutivity principle. The principle states, in rough terms, that if a formula ϕ contains a free occurrence of a subexpression α, and α is identical to β, then substitut- ing β for α in ϕ does not change the truth value of the formula. The principle, generally valid in exten- sional logic, has been long questioned in the context of modal or intensional operators. *believe* and *necessarily* are the most often called for examples in natural language. There is nothing surprising, therefore, in the fact that the principle of substitutivity does not hold in general form in intensional logic. However, Montague (1974e) determined the extend to which the principle can be used in IL. According to this, variables of common type can be substituted for one another without restrictions.

As for other expressions the extensional substitution (like that of *be*) can be used whenever the expression to be substituted for does not stand in the scope of intension operator ˆ. The move assures that extensional part of IL maintains its holdings. In the general case. however. it is required that if two expressions are to be substituted for one another in a formula they must have identical extensions not in one but in **all** possible worlds. In other words for a formula ϕ of IL

- $\phi \equiv \phi_\beta^\alpha$ *only* if $\hat{}\alpha = \hat{}\beta$.

where ϕ_β^α comes from ϕ by substituting all "free" occurrences of α for β. We do not have to add here. that all necessary variable renaming must be taken care of.

Elements of category $T = t/(t/e)$ translate into $(\lambda P (^\vee P(^\wedge\alpha)))$ where α is from WFE_e. Observe that the type e is no longer empty and contains our non-lexical names. In this sense. any proper pseudo-name of category T. like for example *John*. translates into type $<<s.<<s.e>.t>>.t>$. i.e a function from properties of individual concepts P (of type $<s.<<s.e>.t>>$. i.e generalized 1-place relations) into expressions of the form $^\vee P(^\wedge\alpha)$ of type t which denote that a generalized object α of type e ($^\wedge\alpha$ is of type $<s. e>$) has that property. For the full presentation of translation component of PTQ the reader is referred to Montague (1974f). Let us see. however. in one simple example. how PTQ works on the fragment of English. Take a rather straightforward sentence

(1) John runs.

In the categorial grammar of the fragment it is assigned the p-marker shown on Figure 2.9. In PTQ. *John* translates into $(\lambda P (^\vee P(^\wedge j)))$ where $j \in WFE_e$. and *runs* translates into **run** $\in Con_{<<s.e>.t>}$. According to the functional application rule discussed above. *runs(John)* $\in E_t$ (before surface form refinement) translates into $(\lambda P (^\vee P(^\wedge j)))(^\wedge\textbf{run}) = (^\vee(^\wedge\textbf{run}(^\wedge j))) = \textbf{run}(^\wedge j) = \textbf{run}_*(j)$. This means that the property of running is among properties of **individual concept** (or generalized name) of j. In any particular context $<w. t>$ it will denote that an individual j runs at w in moment t. On the basis of Postulate 3 it reduces however to the fact that the individual j belongs to the set of individuals denoted by **run**.

t (John runs)

1/(1/e)   1/e
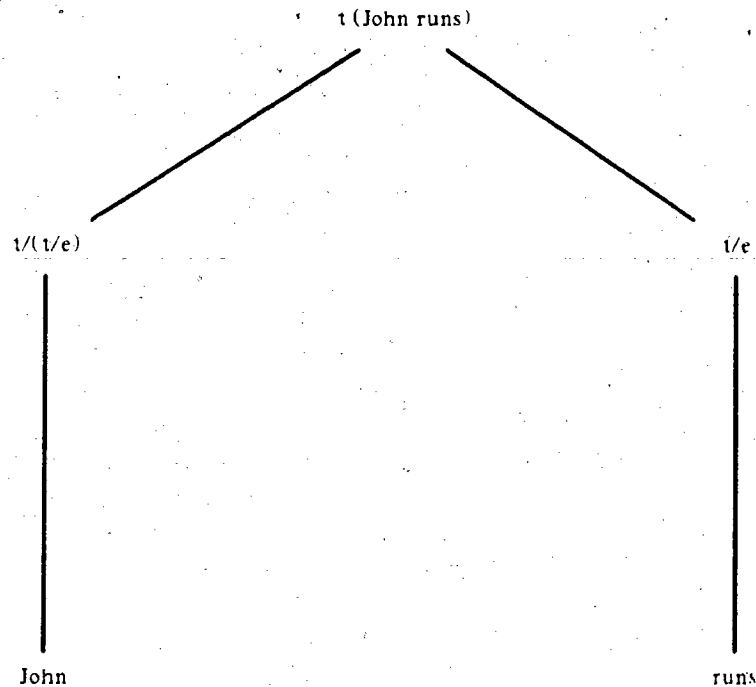
John   runs

Figure 2.9. The PTQ's categorial grammar analysis for 'John runs'

2.1.6. Problems with Possible World Semantics

Possible world semantics is a honest. formal theory. and its most well-known manifestation. Montague's PTQ. is even more honest and formal. A great deal of the formal elegance of the latter stems from the use of intensional logic as the meaning representation medium. This is IL that gives the Montague's system the ultimate appearance of the mathematical theory. It is also responsible (at least partially) for the level of complexity which makes the theory's direct computational realization out of the question at the moment. Two notions. in particular. defy any computational treatment: the intensionality. and the general deduction in IL. The latter is dramatically complicated. far more than first-order logic deduction which. as we know. is only semi-decidable. see. for example. Warren and Friedman (1982) and Warren (1985). The former creates foundational problems which have been sig-

naled ever since the original publication of Montague works on semantics. The problem that seems to be the most significant is inseparably connected to the IL's notion of logical equivalence. and resulting from it version of the principle of substitutivity. Although Montague's solution with the generalized equivalence notion is intuitively acceptable. the problems begin as soon as one starts to talk of the interpretation of attitudes. In possible world semantics the logical equivalence always involves, by definition. sameness of interpretations. Let ϕ be a valid formula, i.e., $\Box\phi$ has the truth value *true*. It follows that ϕ is true in all possible worlds (Montague 1974ef)†. Let ψ be another valid formula; different from ϕ. In IL the two are undistinguishable, as their interpretations (or meanings) are identical. Suppose now that the fact *believe'*$(\alpha, \hat{\phi})$ is true in some set of possible worlds W. Then, by the equivalence of ϕ and ψ the formula *believe'*$(\alpha, \hat{\psi})$ is true also exactly in W. But α (whoever he or she is) may have never heard of ψ! Also, if one day he/she finds out that ψ his/her discovery is reduced to triviality. The problem. discussed to some length by Dowty (1981). included similar kind of criticism of another long-standing assumption that names are rigid designators. More recently the argument was repeated in Barwise & Perry (1983) and Warren (1985). Montague himself was well aware of the problem. although he never took a definite stance toward it. In some places he maintained that we have to accept the conclusion. cf. Montague (1974d). in others he attempted to modify the account so that the logical equivalence has to be replaced by a stronger notion of synonymy in the principle of substitutivity

Let us assume for that B is the interpretation for some language L (called here the "Fregean interpretation"). and $N \times C$ is the set of *points of reference* $<i, j>$ as defined earlier in this section. Let further the subset $A \subset N \times C$ be that of "actualizable" points of reference (we may wish to call them "possible"). Then we define a class K of pairs $(B, <i, j>)$ such that

- $$K = \left\{ (B, <i, j>) \mid <i, j> \in A \right\}$$

---

† Here $\Box$ is understood as the standard necessity operator in normal modal logic. See also Montague (1974be) and Cresswell (1972).

Let now ϕ and ψ be any two expressions of L (we simplify the original Montague proposal for exposi-tory reasons). We say that ϕ and ψ are **K-equivalent** iff both have same denotations in all $<i, j>$'s such that $(B. <i, j>) \in K$. In other words, ϕ and ψ have their intensions identical on A, but not necessarily beyond that. On the other hand, the two are **(weakly) synonymous** iff their extensions coincide over all the set $N \times C$.

What we've done here is to have traded our previous notion of logical equivalence for that of $K-equivalence$. Clearly, if ϕ and ψ are synonymous they are also K-equivalent, but not conversely. In this way the intuitive interpretation of attitudes is being saved. Now, to make a substitution of ϕ for ψ we need both expressions to be synonymous, not merely K-equivalent. That is because K-equivalence guarantees the coincidence of extensions only at these points of reference that are elements of A, while the synonymy requires also all other points of reference, let us call them "impossible", to be accounted for, see also Cresswell (1970 and 1972). Warren (1985), for instance, points out that the meanings of 1+1=2 and the Goedel's second theorem are indistinguishable in the Montague system. But if we allow "impossible worlds" to enter into the image, then the IL notion of necessity becomes very strong, and one cannot tell with all certainity that the two facts above are indeed *necessarily* true. We can say that they are equivalent as far as we can go, but this is not enough, indeed, for IL's logical equivalence. This raises the obvious question about practical applicability of possible world semantics. It appears extremely unclear how one could construct a suitable set of possible worlds for a reasonable fragment of English. When we start from the other end and assume some structure of possible worlds then when interpreting language expressions in such a system we must accept all consequences of this assumption: including a particular instantiation of logical equivalence. The prob-lem does not lie in the theory itself but in developing any practical model to apply it to. This means that although we have managed to represent natural language in formal terms, we still have no means of properly representing its denotational base. We shall discuss these problems more closely in the fol-lowing chapter.

The only other alternative is to reword the theory in more discrete terms, that is, to take advantage of the fact that a conscious individual perceives only a relatively small part of the surrounding reality, and uses his language accordingly. The fact that we understand each other seems to imply that we perceive the reality in the same or similar ways. In this sense, whatever we say or do has a relativistic character, and can only rarely get a clear label as being true or false, good or bad, right or wrong. In such an account a good deal of what is meant by an utterance belongs in ourselves, and an attempt to discover a universal meaning often makes no sense. That is the drive taken by another important, contemporary semantic theory which is said to hold a new promise for a general break-through in the philosophy of language. We refer here to Barwise and Perry's Situation Semantics which will be discussed next.

2.2. Situation Semantics

In part. Situation Semantics evolved out of the frustration that surrounded the development of possible world semantics. especially as enunciated by such theorists as Carnap (1947). Lewis (1976). Cresswell (1973). and Montague (1974a-f). In their world. the meanings of even the simplest statements became sometimes irritating complexes of functions from functions to functions to functions on functions.... Critics complained that the semantic theory somehow lost its "innocence". that it took odds with reality. with the way people perceive and understand the world. It is not surprising therefore that the critics of the possible world semantics. in their search for an alternative. turned to preFregean theories. Indeed. as Barwise and Perry point out in their *Situations and Attitudes*. the troubles with the possible world semantics can be traced back to the Fregean notion of *sense*. The introduction of the third realm of senses (beside the "reality". first realm. and mental states. second realm) was dictated by the failure of the substitutivity principle according to which a substitution of one of logically equivalent elements for another does not change the meaning of the whole. This problem. that is by no means limited to natural language usage. has been most often mentioned in connection with the expression of attitudes. Some of the classical examples of the apparent failure of this principle are inextricably intertwined with the semantics of beliefs. Thus

(1)   *George IV believed that Scott wrote Waverly.*

Scott was the author of Ivanhoe.

So. George IV believed that the author of Ivanhoe wrote Waverly

To prevent this. and similar kinds of reasoning. Frege decided that expressions embedded by *propositional attitudes* like *believe*. do not have their usual references. but instead they refer to *senses* i.e. functions from possible worlds to (usual) references in these worlds. This move abandoned the "innocence" of Fregean semantic theory. and pave the way for the concept of intension.

Situation Semantics is intended to restore the pre-Fregean "innocence" of semantics in which. once again. the interpretation of every language expression has significance in the first realm. this time

interpreted as the realm of situations. Situations are handy to manipulate. They are relatively small and they can pick up almost any arbitrary fragment of reality. In the sharp contrast with possible worlds, the situation theory appears to promote computational realization. Is then Situation Semantics the long pursued "philosopher's stone" for Artificial Intelligence research? Is it an alternative to possible world semantics? Surprisingly, answers to these questions are carefully non-specific. When the first excitement surrounding the emergence of the theory of situations abated, objections began to arise whether the theory is not just an informal "notational variant" of the possible world semantics. Situation theory lacks the formal foundations of Montague semantics, and it carefully avoids any direct reference to lambdas, intensions, or possible worlds. Yet, the theory maintains notions which, although buried in the informal notational system, closely resemble that of set, function, λ-abstraction, and intension, which takes us close indeed to possible worlds.

In the remainder of this section we shall briefly discuss some major aspects of the Situation Semantics as presented by Barwise and Perry (1983). We shall not make any attempt to claim superiority of either Montague semantics or Situation Semantics over the other. In the following chapter we will relate both theories to the Stratified Model of meaning representation, and observe how much of each of them can be utilized in this computationally oriented theory.

2.2.1. Primitive Notions

Although the concept of situation is the key notion of the theory, situations are not primitive concepts. They are rather complex objects built out of individuals, relations, and locations. Individuals correspond to what we commonly consider as singular objects: they are not atomic, but they maintain uniformity across some distance in 4-dimensional space. Relations tie one or more objects (individuals, locations, situations) into some mutual dependency. Barwise and Perry (1983) carefully avoid identifying relations with set-theoretic "sets of pairs": they are primitive notions per se. Relations may be 0-ary (situational states), 1-ary (properties), binary, etc. Locations are defined as connected 4-dimensional regions in the space-time continuum. They may be selected quite arbitrarily, much as

situations can snapshot an arbitrary part of reality. Locations remain in numerous relationships to one another, such as temporal precedence or overlapping, spatial inclusion, disjunction, etc. Beside these primitives, the set-theoretic notion of set is used among basic concepts of the theory. As a consequence, the assumption of an underlying set theory is being made (p. 52).

2.2.2. Situations

"Reality" consists of real situations. The latter are somehow individuated parts of the reality which is a real situation itself, a global one. Real situations may overlap one another and stand in numerous relationships. They may be bigger or smaller, but it is not clear what is required to be in even the most elementary *real* situation. It is an equally open question when one can say that a real situation contains this and that and nothing more, and it begins and ends here and there. It is extremely difficult, if not impossible, to decide whether some aspect of the real world belongs to some real situation, or it falls aside. Consider the simple question of relevance. Many aspects of the reality may seem irrelevant to a situation we are concerned with, and thus may be classified as external to the situation. However, some of these aspects can suddenly become relevant and their role must be acknowledged if the situation is to be complete. Therefore, contrary to our initial assumption they have to be considered elements of the situation. Perhaps we can't even talk of other real situations except the reality itself. If the only real situations were the global ones, we would find ourselves forced to admit that we did not manage to escape the possible world paradigm. This argumentation appears to be one of the most serious problems of the semantics based on situations.

Fortunately we can invent *abstract situations* which are no longer parts of reality, and we can arbitrarily constrain them. Abstract situations may not correspond to their real counterparts. Barwise and Perry introduce the term *actual situation* to refer to those abstract situations which *exactly correspond* to some real ones. This notion is controversial for the reasons explicated above. Thus we are left with *factual* abstract situations, that is, situations that say something about selected aspects of the reality, and are silent about all others. These factual situations create the core of the theory,

Factual situations are useful creations: they appeal directly to our intuition. It seems that all our contact with the surrounding world occurs in terms of factual situations. When we talk, believe, or disagree, we formulate our thoughts with such incomplete abstract situations, which may or may not prove factual, and then reason and act accordingly. In this sense our grasp of reality is somehow fragmentary, our reasoning simplified. We do make mistakes: we are condemned to them. But we still take "chances" because the reality is too complex to be comprehended directly. Factual situations offer some help. Our ability to observe momentarily only the most relevant aspects of the world, and abstract from those irrelevant aspects establishes foundations for our comprehension of "real situations".

Factual situations cannot be ignored, however: the concept is by no means a new one. The utterance of even the simplest sentence as, for example. *John is sitting*, creates an abstract situation which may prove factual or non-factual depending on when, where, and of whom it is claimed, but almost certainly the situation will not be actual. Possible world semantics tries to be actual. When interpreting even the simplest fact possible world semantics accounts for what is, would be, has been, etc going on in the world. Perhaps the most explicit exposition of this approach are Lewis' (1976) index coordinates. This is why we often feel uneasy computationally about this theory, and its full application to AI research looks so hopeless.

Situation Semantics may offer an alternative. But to make it useful one has to find a formal means by which to talk of situations. On the following pages we briefly describe the major aspects of the Theories of Situations and Attitudes, Barwise & Perry (1983). We also comment upon their utility (or alleged utility) to computer applications.

2.2.3. The Theory of Situations

Situations consist of objects: individuals, locations and other situations, and relations which interconnect them. Objects and relations are not unique across situations. In fact they may appear in

many different and remote situations. or as Barwise and Perry call it. they persist. An individual may appear in many different situations. and many unrelated happenings may take place at a specific loca- tion. Relations may persist just as many different people may be sitting at different times and places Also situations may persist as when more that one individuals share some belief or another. More for- mally. we may view abstract situations as being sets of triples $<l. r. i>$ such that l is a location. r is a relation tuple. and i is the polarity value. either 0 or 1. Thus the situation e in which John walks at some location l_1. and talks at some possibly different location l_2 is represented in Situation Semantics as

$$(2) \quad e = \left\{ <l_1 (walks\ John). 1> ; <l_2 (talks\ John). 1> \right\}$$

Barwise and Perry call these constructs *courses of events* or *coe's* (we will often stick to the more universal term of abstract situation). In addition a number of constraints can be added to a *coe* to impose some space-time dependency between participating events. For example. if John's walking temporally precedes his talking we would use $(l_1 \ll l_2)$ in the above coe. This information is also intro- duced as a triple $<l_u (< l_1. l_2). 1>$. where l_u represents the *universal location* such that it contains all other locations. In general l_u will. therefore. denote the entire space-time continuum

A coe at which all relations hold at the same location are called *states of affairs*. These are static situations in the sense that they address a fixed space-time region. in which all relations in ques- tion maintain their holdings. When abstracted over the common location. a state of affairs becomes a *situation type* or. as a possible world theorist would say. a (partial) function from locations to states of affairs. Situation theory does not refer to situation types as functions. In a sense. a situation-type may be identified with some abstract object which. in turn. identifies the class of all situations of that type. Situation-types play a similar role to propositions in possible world semantics but they are far more "concrete"- small. finite objects which. it is hoped. can be represented in a computer. A state of affairs is just a special case of a coe. In fact courses of events are nothing less than sets of states of affairs occurring at different locations. It follows. as Barwise and Perry point out. that a coe is. again. a

(partial) function from locations to situation types. Observe that the triples $<l, r, i>$ may be regarded pairs $<l, s>$ where s is a situation type constructed in the way that allows at most one situation type at a location so that. in effect. we are getting a function from locations to situation types. Let W be a global situation. or a world. Then W is a total function from locations to situation types at these locations. Further. let L be some set of locations. a subset of all possible locations. We have then that a total function from L to situation-types is a coe. or in other words $coe = W_L$. the function W with its domain narrowed to L. The definition is intuitively acceptable. A course of events runs through locations and marks them with situation-types. The only artificial element of a coe is (if included) a triple of the form

(3) $<l_u, (OP\ l_1 \qquad l_n), 1>$

where l's are locations. OP is a relation name. and l_u is the universal location mentioned earlier. Because every coe is extensional the triple carries no information. as the space-time continuum is already ordered. Its role will be justified. however. when we generalize coe's to event-types.

Situation-types are a simple kind of *event-types*. Event-types derive from coe's in which at least one constant has been replaced by a variable. or *indeterminate*. Thus

(4) $\mathbf{r}_{w,t} = \left| <l_1, (walks\ x), 1> \quad <l_2, (talks\ x), 1> \right|$

is the event-type of *walking at some location l_1 and talking at some location l_2*. Barwise and Perry introduce the notion of *role* to stand for binding variable in event-types. To use their notation. the role

(5) $<\lambda_{w,t}, e_{w,t}>$

is that of being the one who walks at l_1 and talks at l_2. The process of deriving event-types is then very much like λ-abstracting in functional analysis. However. Barwise and Perry prefer not to regard event-types as functions. This is a technical detail we should not be concerned with here.

A role may be used outside the situation it derives from, but this situation is carried along with the role. For example, the situation-type

(6) $e_e = \left\{ <l_3, (eats\ x_{e_{wt}}),\ 1>,\ l_2 < l_3 \right\}$

where $<$ stands for temporal precedence

denotes the situations in which someone who walked at l_1 and talked at l_2 is eating at l_3, and talking wholly temporally precedes eating.

Roles, and basic indeterminates, may get bound, or *anchored*, by providing values for them: individuals, relations, and locations, and other complex objects including other roles. Because event-types are not regarded as functions in Situation Semantics, the process of anchoring is not a functional application. Instead Barwise and Perry introduce a special function called *anchor* from indeterminates (and roles) to objects that binds free variables in event-types. The problem with this definition of an anchor is that it has nothing to do with event-type structure whatsoever. A poorly chosen anchor may result in a senseless coe unless the anchor's binding abilities are somehow constrained. Indeed, Barwise and Perry constrain the anchor binding to the effect that it reduces to the process of instantiation of event-types seen as functions from objects to coe's, or other event-types

Despite all these difficulties, the concept of event-type and role appeal to our intuition So far they appear to be just syntactic creations: we still do not know how to interpret or manipulate them. If the language of abstract situations is to become a meaning representation medium we have to establish its semantics. Failing that the theory can hardly be used for any practical (read computer) applications.

For a greater detail of the theory of situations, the interested reader is referred to Barwise & Perry (1983) We now discuss the semantics of situations, and how the theory explains and represents one of the most difficult semantic problems: attitudes.

2.2.4. Meanings

The meaning of a sentence ϕ is a relation $d.c[\phi]e$ between discourse situations d, speaker connections c, and described situations e. *Discourse situation* is the one in which the utterance of ϕ takes place. All discourse situations are classified by a common event-type:

(7) $\quad DU = \left\{ <l, (speaking\ a), 1>, \ <l, (addressing\ a\ b), 1>, \ <l, (saying\ a\ \alpha), 1> \right\}$

where l, a, b and α are indeterminates.

This structure can be further augmented by other elements crucial for identifying an utterance, including the referent of every name or pronoun used in the sentence. Not only names and pronouns need extensions to make an utterance meaningful; every word in the sentence used in d gets instantiated into the described situation e. This process, which resembles that of taking the extension in Montague semantics, is called *speaker's connections*, and is said to be a function from the sentence's words into a described situation. Finally the *described situation* e is the one the utterance of ϕ is communicating about. This is the very situation from which the utterance gets its interpretation.

Situation Semantics preserves the Fregean principle of compositionality of meaning. The meaning of a sentence constituent α is therefore the relation $d.c[\alpha]\sigma.e$, where σ, called *setting*, is provided by other parts of the sentence describing some situation e. In some sense, therefore, a sentence constituent describes an event-type with indeterminates to be anchored by the rest of the sentence. For example, the meaning of the tensed verb phrase *walks* is the relation $d.c[walks]e_x$ where e_x is the event-type.

(8) $\quad w_\lambda = \left\{ <l, (walks\ x), 1> \right\}$

So, the meaning of a sentence constituent may be identified with an event-type thought of as representing a class of situations of that type. Here the event-type w_x describes the property, a one-place relation, of walking. This should give us a better idea of what the properties, and relations are in Situation Semantics.

The same analysis applies to the meanings of singular noun phrases with the exception of definite descriptions. Here a new element comes into the image: *resource situation*. In order to identify the unique object determined by a definite description the speaker (as well as the addressee) must exploit a resource situation satisfying the definite condition. This situation, which may often be different than the described situation, identifies the referent of the definite description so that it can be uniquely referred to in the described situation. If we say

(9)   *The man who is talking to John is a fool.*

the resource situation in which there is a unique man presently talking to John is obviously a part of the described situation. But in the utterance of

(10)   *The man whom we met in library is a fool.*

the resource situation may be quite remote to the one we describe. but certainly both must contain the individual in question. The role of the resource situation in establishing the meaning of definite noun phrases is very important. It provides the necessary context for a definite description, so that we can talk of referential, or *value-loaded* use of the latter. When, on the other hand, the resource situation is not available at the time of discourse we have a non-referential, or *value-free* use of a definite description. In the latter case what a definite noun phrase describes is a function (or relation) between possible resource situations and individuals.

When we talk of the meanings of common noun phrases, we invariably think of quantifiers. What would be the meaning of such words as *a. the. every.* etc? It is certainly not enough to identify *a man* with the property of being a *man,* or, in other words, with the event-type

(11) $m_u = \left\{ <I. (man\ u). 1> \right\}$

What we want is a relation between the property of being man, and some other, undetermined property Q, see Cooper (1985), Cercone & Schubert (1975). For example, the meaning of *Every man*

walks will be represented † by the relation *every* between the property m_u of being a *man* and that of

walking. w_x. Similarly. in *A unicorn lives in the park* we are getting the relation between properties

of being a *unicorn* and *living in the park*. This is all very much like the ordinary logical account of

quantifiers: when the quantified phrase appears at subject position the other property Q is provided by

the following it verb phrase. What happens. however. when the quantified phrase occurs at some

other position in a sentence? In *John seeks a unicorn* the property of being *unicorn* is related to some

unknown property Q we may try to guess. One apparent candidate for Q is just the property of being

sought by John. and this effect can be reached easily by transforming the sentence into the passive

voice. Passivation parallels the PTQ's *de re* derivation in categorial grammar. and determines the sub-

sequent translation into IL. Nevertheless. there remains the possibility that the property Q is not that

of *being sought by John* because of non-referential character of the utterance. A good guess would be

to identify Q with the property of *being found by John*. but there is nothing there to prevent us from

assuming other possibilities. Some of the other possibilities may be excluded for violating certain

common-sense restrictions. but we are left with a great many options. We can try. of course. to con-

strain the number of alternatives to just one alternative. perhaps an implicit property that derives from

the property of being sought by John. We shall discuss this point in depth in Chapter 4.

We come to one of the most important issues of any semantic theory: the notions of interpreta-

tion and truth. Let us concentrate here on interpretations of utterances of indicative sentences. some-

times called statements. Later we'll briefly discuss how Situation Semantics interprets utterances of

selected types of noun phrases. The contribution of other syntactic categories is not yet sufficiently

elaborated in the theory.

As mentioned. in Situation Semantics the meaning of a sentence ϕ is a certain relation $d.c[\phi]e$

between discourse situations. speaker's connections. and described situations. In this way the meaning

of ϕ is independent of any particular utterance. and no commitment is made about truth or falsity of

---

† To be more accurate we should say that both properties of manhood and walking. as well as the relation between them
are constituents of interpretation of (every) utterance of the sentence. In other words. they are parts of every situation e such
that $d.c[every\ man\ walks]e$.

ϕ. We are restricted, however, in the way in which we can interpret any particular utterance of ϕ: the meaning *underdetermines* the interpretation, to use Barwise and Perry's words. Sentences do not have interpretations, their utterances do. When we utter a sentence ϕ at any particular occasion it will uniquely establish the discourse situation d and connections c. But there may be any number of situations that the utterance can describe, including 0. The *interpretation* of an utterance u of a sentence ϕ is therefore the set of all situations e such that $d.c[\phi]e$ holds † More formally, let us use the symbol $I_{d,c}[\phi]$ for the interpretation of an utterance of ϕ at d with c, then

$$(12) \quad I_{d,c}[\phi] = \left\{ e \mid d.c[\phi]e \right\}$$

Therefore, a sentence may be identified with the class of all its utterances past, present, and future. The meaning of a sentence uniquely determines the set of all situations that may ever be described by any utterance of the sentence. Some of these situations will be included in other, bigger situations, so there results a structure of situations at the bottom of which there are coe's that contain just the facts expressed in any particular utterance. In this way the assignment of meaning to sentences is more accurate than in possible world semantics. Observe that sentences like α and α & $(\beta \cup \neg\beta)$ will have different meanings in Situation Semantics. When we identify meanings with sets of possible described situations then $M(\alpha \& (\beta \cup \neg\beta)) \subset M(\alpha)$, because a minimal coe described by an utterance of α will be smaller than that described by an utterance of α & $(\beta \cup \neg\beta)$.

Having the notion of interpretation for statements we can move on to define what it means for an utterance to be true. For this purpose Barwise and Perry introduce the notion of persistence. A statement is said to be *persistent* if whenever it describes a situation e, it also describes every situation e_1 such that $e \subseteq e_1$. It can be easily seen that e_1 extends to a world. It follows that if a persistent statement is true in some situation e it must also be true in any extended situation e_1 which includes e. Barwise and Perry give the following definition of truth for persistent statements. A persistent

---

† In other words, all situations in the interpretation of some utterance contain some smallest subsituation (a coe) which is explicitly described by this utterance. See, however, further pages for more discussion.

statement is *absolutely true* if there is an actual situation e_a in its interpretation. In other words:

(13) $\phi_{d,c}$ is absolutely true iff

$\exists e_a, e_a$ actual, such that $\acute{e}_a \in I_{d,c}(\phi)$

This definition resembles that of satisfiability in the possible world semantics, especially when we assume that the only actual situations are the worlds. Recall that a sentence was satisfiable in Montague's (1974f) definition if, for a given index value $<w, t>$, and a value assignment coordinate g, its extension $Ext_{U_g}^{w,i} = 1$. Here $w=e_a$, and t and g are provided by the discourse situation and the speaker's connections. The notion of truth in Situation Semantics is weaker than that in Montague's system because Situation Semantics deals with utterances of sentences, while Montague treats sentences themselves.

When it comes to the interpretation of (singular) noun phrases we must take into account the assumption of compositionality of meaning. Thus whatever happens to be the interpretation of a noun phrase it must be a constituent of interpretation of any utterance containing it. Let α be a noun phrase, and $I(\alpha)$ be its interpretation. Let $\phi(\alpha)$ be a sentence containing α, and $\phi_{d,c}(\alpha)$ be an utterance of $\phi(\alpha)$. Then

(14) $\forall \phi_{d,c}(\alpha) \, \forall e. \ e \in I_{d,c}(\phi) \rightarrow I(\alpha) \in e$

What we are getting is a function from situations to some of their parts. Or, to reword it in Montague's terms, a function from possible worlds to subsets of these worlds. In the possible world semantics the function is taken as the meaning of the noun phrase. The interpretation is evaluated at any particular world, and with some specific value of assignment coordinate. This same move is made in Situation Semantics. The interpretations are elements of situations. For value-loaded use of α the interpretation will be an individual. For value-free use it will be a function from situations to individuals; this function is the constituent of interpretation of every utterance u containing α.

Our account of meaning in Situation Semantics would not be complete if we did not mention how one can deal with non-persistent statements. In everyday life we utter things which are not

persistent. but still we manage to classify them as either true or false (if they make any sense. of course). Sometimes we use non-persistent statements. not concerned about wider consequences of such utterances. *No one is walking on the road* may be true here and now. but it is certainly non-persistent. On another occasion we can produce non-persistent statements quite deliberately. In *Everybody is here. let's begin* the speaker's attention is concentrated entirely on a selected group of people in some auditorium. those who are to play active roles in a prepared experiment. or so on. Certainly not *everybody* is here. probably not even every person who is supposed to attend. In this sense the statement is not persistent. It is not difficult to see that many utterances asserting uniqueness. universal quantification. enumeration. stating proportions. etc. will be often non-persistent. Singular definite descriptions are almost inherently non-persistent. To account for meanings of non-persistent statements Barwise and Perry introduce the notion of truth relative to some actual situation *e*. and say that a statement *u* of ϕ is *true relative* to an actual situation *e* if $e \in I_{d,c}(\phi)$. Now our earlier notion of absolute truth for statements no longer suffices. For a statement to be (absolutely) true it is not enough to have an actual (factual) situation in its interpretation. We must account for the statement in a wider context. What is required is some locally largest actual situation e_u say the one the statement in question is referring to. The statement may not describe this situation fully. but the situation somehow identifies the largest context we are concerned with. Then our statement *u* of ϕ will be absolutely true if $e_u \in I_{d,c}(\phi)$. If we are skeptical about the actual situations. we can say that *u* is absolutely true in a world *w* if $w \in I_{d,c}(\phi)$. This last definition. however. appears much stronger

The notion of non-persistence is not present in possible world semantics. But it does not mean that the theory cannot treat such cases. Indeed. the concept appears dispensable and every statement can be thought of as persistent if only its interpretation is properly constrained. If one says that no one is walking on the road. he means (unless he is making a global assertion which may be false) that no one is walking on some particular part of the road. Similarly. we can make clear that by *Everybody is here. let's begin* we mean actually *every one whom we need. now*. Constraining interpretations of statements in possible world semantics is accomplished by the pragmatic stratum of the theory. and

this approach is very convincing. Some statements will be insensitive to the pragmatic context. and these we can identify with persistent ones. Others which undergo some pragmatic evaluation before we can interpret them fully may be called non-persistent. So. in possible world semantics all statements are persistent. The only problem with the pragmatic (contextual) coordinate is that one never knows how much pragmatic information is necessary to make a statement persistent. But this aspect appears to fare no better in Situation Semantics: the situation e_u will always be relative to the statement we make.

To this point we summarize the major differences between Situation Semantics and Montague's approach. As we observed at the end of the last section. Montague developed a very powerful system for representing meaning of natural language expressions. but he virtually neglected the problem of representing the language's denotational base. Except for rough intuitions. we know nothing of possible worlds and far less how to construct them for any practical use. The theory of situations takes the opposite approach trying first to build a discrete model of the "reality". then suggesting ways to relate our language to this model. Situation theory retreats from some conclusions made in possible world semantics. e.g logical equivalence. necessity. generalized notion of truth. but. as we saw. others remain virtually unchanged. One of the most significant consequences of situation theory is the locality of the notion of truth. Because Situation Semantics evaluates utterances and not sentences. some syntactic manipulations on the language expressions may no longer be uniformly reflected on the semantic ground. In particular. the intensional logic's principle of substitutivity has to be replaced by a collection of weaker special cases. and its applicability carefully constrained. We shall discuss these. and others problems. briefly in the following subsection.

2.2.5. Attitudes

A typical attitude report construction in language consists of an individual. a transitive attitude report verb. and a sentence or clause describing the situation being reported. Thus

(15) *1. John saw the temperature be 100.*

    2. Mary saw John eat the cookie.

    3. Mary saw that John ate the cookie.

    4. John knows that Cicero was a famous Roman orator.

    5. John believes that a unicorn lives in the park.

    6. Bill doubts that a unicorn lives in the park.

are all classified as attitude reports. The above definition is purely syntactic. and as such insufficient to

properly capture the characteristics of the attitude report phenomenon. There is an important semantic

feature which distinguishes the "true" attitude reports of (3) to (6) from a mere descriptions of some

perceptual events as (1) and (2). The former. called *epistemically positive*, involve some mental par-

ticipation of the individual. while the latter, called *epistemically neutral*, do not assume such partici-

pation. Indeed, the epistemically neutral reports cannot give a basis for concluding any mental or emo-

tional participation of the agent: That is. from the fact that *Mary saw John eat the cookie*, one can-

not tell anything about the state of Mary's mind. One cannot even infer that *Mary saw that John ate*

the cookie. as she may say that she did not know it was the cookie or that John was eating it. etc

So. epistemically neutral reports do not report attitudes. at least not attitudes of the agent. Mary can-

not be said to have any attitude toward John's eating the cookie: she may have even not recorded the

fact. It is the speaker of (2) who may have it. On the basis of (2) the speaker may come to believe. or

doubt. or something about Mary's emotional state In this sense the epistemically neutral reports are

quite unlike those epistemically positive. This can be seen more clearly in connection with some logical

principles we discuss below. But the distinction remains interesting: Epistemically neutral reports do

not cause any of the logical problems that are so characteristic of "true" attitudes. so we can narrow

our investigation space to the latter class. The same approach is taken toward remaining attitudes.

and it is consistent with the discrete character of the semantic theory. Because we cannot generalize

beyond our discrete model. we have to trace down the sources of every relevant phenomenon. approxi-

mate its scope and state well defined exceptions to rules. There remains however. a class of inherently

logically problematic attitude report constructions which create the core of the phenomenon discussed for years in philosophy and logic. We shall turn our attention to the Situation Semantic treatment of these.

The special place attitude reports have occupied in semantics and philosophy of language is. at least partially. due to their apparent failure to obey some of the most fundamental logical principles which were long considered at the core of mathematical logic. These are (as listed by Barwise and Perry) the principles of veridicality. substitution of logically equivalent elements. existential generaliza- tion. negation scoping. distribution of conjunction and disjunction. It is not true that the principles fail for any instance of attitude report. or even that the failure range is the same for every attitude. What worried philosophers and logicians are the versions of the above principles we state below. We assume for that that **att** is an arbitrary (true) attitude report construction like α *believes that.* α *knows that.* or α *sees that.* and ϕ is an embedded sentence or phrase being reported. Occasionally we may use a parameter with ϕ for example $\phi(t)$. to explicate that t is a constituent of ϕ.

1. **Veridicality Principle**

    If **att**(ϕ) *then* ϕ. $\square$

2. **Substitution of Equivalent Elements Principle**

    If **att**$(\phi(t_1))$ *and* $t_1 \equiv t_2$ *then* **att**$(\phi(t_2))$

    where $\equiv$ is the logical equivalence. $\square$

3. **Existential Generalization Principle**

    If **att**$(\phi(the\ u))$ *then* $\exists x\ \phi(x)$

    If **att**$(\phi(a\ u))$ *then* $\exists x\ (u(x)\ \&\ \phi(x))$. $\square$

4. **Negation Scoping Principle**

    If **att**$(\neg\phi)$ *then* $\neg$**att**(ϕ). $\square$

5. Conjunction and Disjunction Distribution Principle

If att($\phi\&\psi$) *then* att(ϕ) & att(ψ)

If att($\phi\cup\psi$) *then* att(ϕ) $\cup$ att(ψ). $\square$

The invalidity of the Veridicality Principle for the attitudes like *believe*, *doubt*, or even *assert* has been long acknowledged. From the fact that John believes that a unicorn lives in the park we cannot conclude that a unicorn lives in the park. On the other hand, for attitudes like *know* or *see* the principle holds all right. One cannot know ϕ without ϕ being the case. Similarly one cannot *see* or *see that* ϕ without ϕ being the case. This is quite an important conclusion. If you say *I saw that the tempera-ture was 110*, being unconscious victim of the paralaxa phenomenon, while the true temperature was 108, then your claim was false. You did not see the temperature to be 110. What you saw, from your position, was the thermometer looking as if it pointed at 110. So if something was not the case you could not see it. You might believe you saw it, of course.

There is nothing disturbing about this principle, except perhaps that it once served as an evidence that in a semantic theory the Fregean principle of compositionality of meaning and the assumption that the meanings of sentences are truth values cannot coexist. The former survived, so the latter had to give the way. Recall, for example, the possible world semantics, generalized truth values.

The failure of the Substitution Principle for attitude reports triggered a long debate in philosophy and logics. Mostly because it formed the foundations of extensional logic reasoning. The solution to the problem, accepted in the possible world semantics, was the redefinition of the notion of logical equivalence. According to this revised definition which laid the foundations for emergence of intensional logic, two expressions were equivalent if and only if their intensions coincided. This revision preserved the validity of the principle of substitutivity and extended it over attitude reports. This, however, gave rise to other problems which we discussed in the last section.

Situation Semantics does not offer any better solution. It states, not without reluctance, that the principle generally fails for non-referential uses of substituents, and is mostly valid in referential cases. The authors' observations, although still too informal and incomplete to be taken as an evidence, nevertheless, are quite convincing. They replaced the intensional logic equivalence notion stating the so-called *Weak Substitution Principle* which is said to work for all non-negative attitudes.

6 Weak Substitution Principle

If t_1 and t_2 are both either referential or non-referential *then*

if $\mathbf{att}(\phi(t_1))$ and $\mathbf{att}(t_1 = t_2)$ *then* $\mathbf{att}(\phi(t_2))$. □

The principle holds for all attitude report verbs discussed in the book, with the exception of *doubt that* considered as lack of belief. This exception is quite apparent, and by no means surprising. If $\neg\mathbf{believe}(\phi(t_1))$ and $\neg\mathbf{believe}(t_1 \equiv t_2)$ then one certainly cannot conclude that $\neg\mathbf{believe}(\phi(t_2))$. If John does not believe that Tully was a famous Roman orator, and he does not believe that Tully was Cicero, he may still believe, or even know, that Cicero was the famous Roman orator. Unfortunately, so stated the principle is plainly invalid. It seems acceptable for simple attitude report cases, but fails whenever an iterated attitude is involved. Suppose for that ϕ stands for *Mary believes that t_1 was a famous Roman orator*, and t_1 and t_2 are *Cicero* and *Tully* respectively. Let then *att* be *John believes that*. Then from the facts that John believes that Mary believes that Cicero was famous Roman orator, and that John believes that Cicero was Tully, it follows by the above principle, that John believes that Mary believes that Tully was a famous Roman orator, an obviously wrong conclusion. An even more shattering counterexample can be constructed with the verb *know* for which the Veridicality Principle holds. For non-iterated cases of (positive) attitude reports the Weak Substitution Principle seems valid but it does not get us any closer to a better understanding of attitudes.

Existential Generalization Principle, known also as the existential quantification scoping problem, has its source in a wider linguistic phenomenon of distinguishing between referential and non-referential uses of certain classes of noun phrases. In general the principle is valid for referentially used

noun phrases, while invalid for the others. For attitudes the principle is a combination of the veridical-ity principle and the quantifier scoping problem. We shall discuss the issue in Chapter 4.

The principle of negation scoping is said not to work for beliefs, doubts, or asserting, although, as Barwise and Perry point out, we are willing to accept it in certain cases. If I believe that Jack is not sick, then I certainly do not believe that he is. The only cases where the principle fails appear to be confusions with the principle of substitution. I may believe that Jack is not sick, but at the same time I can strongly believe that the person I saw from a distance last night, not recognizing him as being Jack, is apparently sick. Barwise and Perry conclude that if the negation scoping principle were valid for beliefs we would have to accept the contradictory information that I believe and do not believe, at the same location, that Jack is sick. This confusion comes from the implicit use of substitution princi-ple which, as we saw earlier, does not hold for attitudes. There is nothing contradictory in believing that Jack is not sick, and believing that the guy I saw last night is sick. It might, after all, happen that the guy was not Jack, as I certainly assume, or rather do not assume he was

The principle of conjunction distribution holds for positive attitudes that is those that do not involve negation. The parallel principle for disjunction generally fails. Exactly the opposite happens for negative attitudes as *doubt that* by virtue of some elementary logical laws. The principle of disjunc-tion distribution works for epistemically neutral *see* (called see_n in the book) for quite obvious reasons as the evidence for that kind of report is given entirely by the surrounding situation, and does not involve the agent mental participation. It also works for so-called primary use of epistemically positive *see that* (called see_p). where the agent actually witnesses the situation he is reporting. This comes as the result of the observation that if α $sees_p$ *that* ϕ then he also $sees_n$ ϕ^*, where ϕ^* is the infinitive form of ϕ. Primary *see that* report is therefore not much different than that of see_n, the only difference being in the assumption that in see_p report the agent actually recorded the reported situa-tion, i.e., he actually comes to know that ϕ. but it does not involve any more significant mental pro-cess. The distinction, although interesting, does not help much with other attitudes

Let us now discuss briefly the Situation Semantics approach to the meaning of attitude reports
The meaning of an epistemically positive attitude report att(ϕ) is defined as the relation

(16) $d.c[att(\phi)]e$

between discourse situations speaker connections and described situations e The definition would
not be very informative if it did not relate the meaning of the report to that of the reported event as
required by the Fregean principle of compositionality of meaning Let therefore a be the agent of some
attitude report att(ϕ) We have

(17) $d.c[att(\phi)]e$ iff $\forall e'$ either $d.c[\phi]e'$ or e' is incompatible with e_0

where e_0 is the situation a is aware of by means of his/her attitude att towards it

The notion of incompatibility is relative to what a sees knows believes or asserts etc in e as well
as what he/she is concerned about In a simplified account e' is incompatible with e_0 if $e' \bigcup e_0$ is
inconsistent. i e contains some contradictory facts That means that every situation e' which is not
incompatible with e_0 is to be described by some utterance of ϕ These situations create the class of
appropriate *att alternatives* (seeing alternatives believing alternatives knowing alternatives etc) of
the agent We must of course state some constraints on constructing these alternatives very much
like in the case of constraining "the other property" in *de dicto* readings of noun phrases The most
elementary constraint would be to require that every agent s att alternative has to contain the reported
fact. cf Cooper (1985) In general. this may be too strong a restriction but it gives us a pretty good
idea what the alternatives are Indeed they resemble Montague s intensions Alternatives are possible
situations. and this begins to sound familiar

Therefore. a is performing the following mental process He excludes all situations e' which are
incompatible with e_0. In all the remaining situations ϕ must hold to guarantee that indeed a *atts* ϕ
The interpretation of att(ϕ) given d and c will be then the set of all these situations e such that for
every situation e' either the relation $<l. (incompatible_{att} a e')$. $1>$ is in e or $d.c[\phi]e'$ holds The atti-
tude will be true if at least one of e s is actual (or factual). that is. there exists at least one situation

e_1 such that $d.c[\phi]e_1$ and $<I. (incompatible_{all} \alpha e_1). 1>$ is not in e. and there is no situation e_2 such

that $\neg d.[\phi]e_2$ and $<I. (compatible_{all} \alpha e_2). 1>$ is in e. Ironically. α's attitude can be determined by

the set of sets of situations which are incompatible with what he/she knows (for seeing and knowing).

or knows and believes (for believing and doubting). Knowing is also restricted by general laws. external

to α's mind. If what α "knows" does not agree with reality. his apparent knowledge appears mere

belief

A disturbing thing about this approach is the informality of the notion of *compatibility*. Intui-

tively. a situation e_2 is *compatible_{all}* to a situation e_1 if $e_1 \subseteq e_2$. and e_1 *involves_{all}* e_2. Unfortunately.

the relation of involving is rather subjective. and depends a good deal of somebody's mental state. If

John believes that Bill took his book. then he probably has considered a number of alternative courses

of events that might have resulted in his book not being where he left it last night. After exclusion of

all incompatible situations. such as himself putting the book in another place. John comes to the con-

clusion that all remaining situations include Bill walking away with the book. That is why he believes

that Bill took his book

2.2.6. Conclusions

In the last two sections we discussed often simplifying the account. two contemporary semantic

theories primarily aimed at providing a satisfactory account of semantics of natural language. two

quite different theories. On one side we have the possible world semantics. very formal and elegant

with roots in modern logics from Russell and Frege onward. The rich mathematical apparatus of the

approach. often identified with the Montague's intensional logic system. gave it the characteristics of a

formal mathematical theory. with all the consequences of the model-theoretic foundations. On the

opposite side (but as we saw not at an antipode) there emerged the "innocent". intuitive. and to some

degree informal Situation Semantics which from its first statement breaks with the Fregean concepts.

and looks for its foundations in earlier semantic systems. Although built on some set-theoretic base.

the theory tries to work out its own methodology which. at the time being. is anything but a formal

system Some critics of Situation Semantics maintain that the theory cannot be taken seriously until its methodological foundations are well defined.

For us in Artificial Intelligence there is hardly time to wait until the foundational problems have been settled. If a theory, no matter how informal it happens to be, offers some advantages over the existing approaches, we shall try to utilize it in our work. Situation Semantics does not give answers to many questions that plagued possible world semantics; and the cost of maintaining the "innocence" of the theory may outweigh its alleged virtues, as seen from the computational viewpoint. The power of the theory stems from the ingenious insight in the way people may actually comprehend the world by manipulating, mentally and physically, numerous sorts of situations. The concept of situation is not a new one, but we believe this is the first time it has been made so central and so explicit a notion in a semantic theory. Backed by the attainments of modern philosophy and set theory, Situation Semantics tries to challenge the overly complex, "non-American" Montague system. But when one looks closer at both theories it can be seen that they meet in many places. This is a comforting observation. The theories do not contradict each other, and in fact they may be considered as mutually complementary.

In the following chapter we shall present an outline of an alternative model for natural language understanding. It is not to say that we shall propose a new semantic theory in this thesis. What we suggest is the way how a computationally oriented system should be built. We shall also show how the major semantic theories of language relate to the framework, and why their direct computer applications seem questionable. The proposed model, which we call *A Theory of Stratified Meaning Representation*, has a very high-level character, and we do not intend to convince anybody univocally at this stage that this is the only, or even the best direction. In the following chapter we shall only loosely suggest how the model can get instantiated in any practical situation, and why this approach promises computational tractability. The next three chapters will actually fill selected slots in the framework. Although no attempt is made to build a new semantic theory for natural language, we shall demonstrate the ways in which the most valuable features of the above presented semantic systems can be harmonized giving a ground for a more powerful alternative

Chapter 3

A Theory of Stratified Meaning Representation

3.1. The Stratified Model

Any meaning representation must mediate between a universe (or a world) which is to be described. and a language which is used to express (or represent) whatever one knows, observes, foresees, imagines, remembers, disagrees, believes, etc. about that universe. The language is equipped with syntax and semantics rich enough, to be able precisely to represent everything we want to represent. In mathematical analysis, for example, we use a carefully designed symbolic language to communicate about functions, sets, and their properties and behavior. In every day life, we speak our natural (human) language talking about surrounding reality and non-reality. Both situations can be included in the model outlined above, yet there is some significant difference between them. This difference lies in the "distance" between a meaning representation language and universe in this bipolar model (Figure 3.1). In the case of the symbolic language of mathematical analysis, this distance is rather small, in the sense that no intermediate-level representation is necessary to map the language into the universe that is, to establish a correspondence between well-formed expressions of the language and parts of the universe. This situation complicates significantly when we examine human natural languages as meaning representation languages. Human languages evolved over centuries, and they are so complicated and sophisticated that they require some non-trivial decoding process in order to uncover the mapping between a language and "reality". We can therefore revise our bipolar model as illustrated in Figure 3.2. Now, we can select some artificial, relatively small universe, e.g., a blocks world, or extensional database for example, and devise a language which directly manipulates that

87
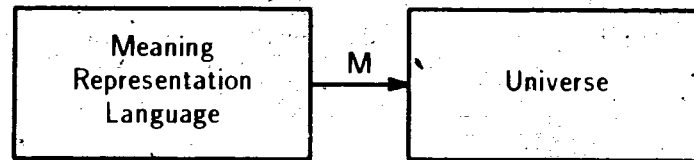
Figure 3.1. The bipolar model of meaning representation

universe. Some translation program would then map a limited subset of natural language into this data

manipulation language. Problems emerge soon after one attempted to enrich the structure of the

universe to make things behave more like the real world. A more complex language is needed to

express properly what is going on in the universe. Subsequently, the language decoding process com

plicates considerably. Before long our translation program segments into a number of not necessarily

sequential subtasks. each of which presents a separate research and programming challenge At the

same time the gap between the ultimate meaning representation level and the universe widened. We

could no longer directly manipulate the universe with the meaning representation language. Most of

the early AI systems fell into this general framework

The denotational base for human languages is so extremely complicated that the design of a

full-size mapping from the decoding level into the universe may be an enormous task. assuming that

the decoding level preserves the expressive power of the source language. Although we may not
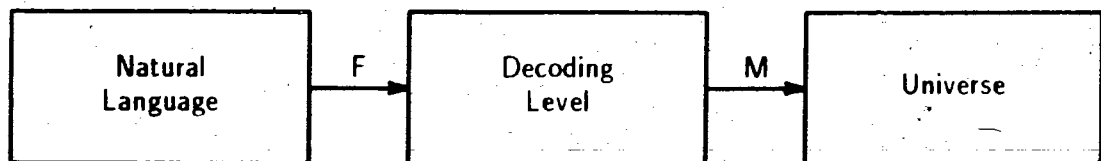
Figure 3.2. A revised meaning representation model for natural languages

comprehend the universe directly. we can maintain an encoded image so that the mapping could be directed into it. Our model of meaning representation. therefore. can be modified once again as shown in Figure 3.3. Thus we set off simultaneously. from both ends of the model. and create two intermediate levels of **Source Language Approximation (SLA)**. and **Universe Approximation (UA)**. † The idea of approximating reality is almost as old as science itself. When the universe did not fully cooperate we often created simplified models to instantiate one observation or another. If we chose our model wisely. we could discover some actual laws. Although these models gave perhaps only a retouched image of the universe. they were easier to manipulate and speculate upon the original world. Yet. if we wanted our discoveries to have any practical significance. we had to provide a means of mapping the model back into the original universe. This process might involve restricting our laws to only well-defined situations. but that depends to a considerable extent on how precise our model happens to be. In general. the more accurate a model. the more credible our findings. For example. Newtonian mechanics once appeared to be a precise approximation of the real world until a relativistic theory showed how simplified a model it was. Nevertheless the Newtonian model can quite appropriately be mapped into a well constrained part of the universe.

Analogously for language translation. we may find it easier to think of the universe approximation as a step by step process. We can build an n-degree approximation which does not have to be done
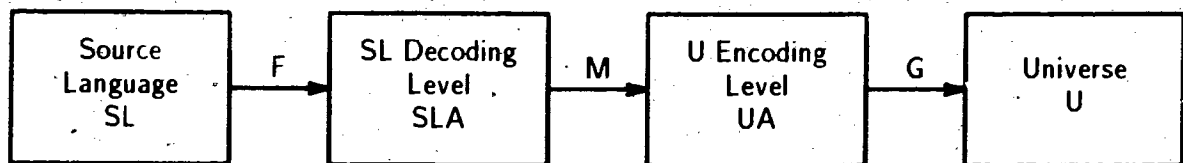
Figure 3.3. Bipolar meaning representation model with bilateral approximation.

---

† The process of deriving the universe approximation is of course from U to UA. even if in Figure 3.3 all arrows lead in one direction. In fact these arcs can be traversed in both directions. and the one we chose here is to stress our main interest in the

sequentially. There is nothing, however, to prevent us from formulating our present theory at a comfortably idealized level. Thus we have a language and its subsequent decoding levels, and there a universe with all these encoding models. How do these both side approximations relate to one another? Well, they might be just two independent processes driving in opposite directions toward some more or less close encounter. In our theory, however, we assume that for every language translation step there is a respective move (backward) from one universe model to another. In other words, every meaning representation level for the language should have an accompanied universe model such that it is fully capable of describing it. In this sense, we reduce the problem of describing the relationship between the source language and the original universe to describing a mapping between n-th degree approximations of both. One may object here by observing that not every step on the Left Side (language) will change the face of the Right Side (universe). If this is the case, we can always introduce some identity transformations where necessary. But, for now, the universe modelling is far less understood than the language manipulation.

Now we put our discussion into more formal terms, and add our intuition about what underlies this general framework. Let $F: SL \rightarrow SLA$ and $G: UA \rightarrow U$ be the language decoding transformation, and the universe encoding transformation, respectively. If F and G can be defined then the original problem of discovering the mapping between SL and U will have been reduced to (probably) a somewhat easier to formalise relation between SLA and UA. There are two major factors influencing the relative difficulty of this mapping we want to reduce. One is the inherent ambiguity of source language. The other comes from the lack of a sufficient fragmentation of the universe in the sense that we could always select a piece of it which *exactly* corresponds to a particular language utterance. This is, among others, what the Stratified Model is supposed to resolve. Yet, we have no guarantee that the one-level approximation depicted in Figure 3.3 would suffice. In general, one can imagine a sequence of subsequent (but not necessarily sequential) both-side reductions $F_1, \ldots F_n$, and $G_n, \ldots G_1$,

---

language processing problems.

such that $F_i : SLA_{i-1} \rightarrow SLA_i$, and $G_i : UA_i \rightarrow UA_{i-1}$,† and such that the bipolar model $\langle SLA_n, UA_n \rangle$ needs no further reductions. Ultimately we obtain a stratified mapping from SL into U, as shown schematically below.

(1) $SL \rightarrow_{F_1} SLA_1 \rightarrow_{F_2} \cdots \rightarrow_{F_n} SLA_n \rightarrow_M UA_n \rightarrow_{G_n} UA_{n-1} \rightarrow_{G_{n-1}} \cdots \rightarrow_{G_1} U$

where M is a mapping in the bipolar model

The model presented above is quite general. and idealized to some degree; we shall be a little more specific about it later. The first point to be made is that the structure presented in (1) (slightly simplified for expository reasons) can be considered as an 'ever-busy' machinery. possessed by some intelligent individual. which passes information through and back between SL and U. Part of this flow will accumulate at some SLA_i levels. preferably at these levels closer to the centre of the structure. This information will then provide a necessary context. knowledge. beliefs. etc. for interpreting further flow. In this sense the transformations on both sides can be steadily enriched. Another alternative is to relegate the knowledge and beliefs to some external knowledge base. thus making the model user-independent In such a case. the structure in (1) is incomplete because. to make it work. one has to bring in some individual knowledge base and plug it into a proper place in the model. By doing so we can guarantee that the transformations will get the necessary support. Although we do not favor either of these possibilities. for the current presentation we assume the latter interpretation When talking of transformations F_i's and G_i's. we shall always assume that an adequate knowledge base is being provided.

The second question which arises quite naturally is how many strata are there. or perhaps. what is the minimal number of them in a realistic natural language system. It is not easy to state any concrete numbers but the need for at least one level $\langle SLA_1. UA_1 \rangle$ over the original language and universe should be acknowledged (see our discussion earlier in this section). We feel. however. that a few more steps will be necessary. and. most important of all. they cannot be reduced to the single

---

† It should be relatively clear that transformations F_i (and G_i) are not functions. in general. therefore the notation $F_i \subseteq SLA_{i-1} \times SLA_i$ may seem more appropriate at times.

stratum scheme of Figure 3.3 without a serious damage to the model. Both the language and the universe can be ambiguous in many ways. and a practical model has to define explicitly all transforma- tions required.to reach the ultimate stratum. A hint in this matter may be provided by the fact that at different stages of language processing we require different amounts and quality of extra-lingual infor- mation for interpreting language expressions. Too much external knowledge. especially when applied at a too early stage. can be as undesirable as its lack. Recall for that the difficulties faced by the so- called "semantic grammar" approach to natural language understanding. The question is difficult to answer also because of the lack of balance in our knowledge on what to expect on the Left Side (the language matters) of the structure and its Right Side (the universe models). While we know a good deal of how to process the language: the universe representation issues were rarely addressed beyond a one degree approximation. The theory of situations discussed in the last chapter may shed some new light on these problems. and we shall discuss the relationship of this theory to the Stratified Model later in this chapter. For the moment. we concentrate mostly on the Left Side. as this thesis addresses the problems that we consider to belong somewhere to the left of the M transformation. However. some issues like ambiguity for example. are characteristic to both sides.

There are many ways to model the Left Side. if we forget the accompanying Right Side for a moment. Specific instantiations may vary. but from the current AI practice one can expect something like the following. Some early F_i transformations (for $i = 1.2.3$) would be concerned with such linguistic problems like phonology. lexical analysis. or syntactic parsing. Those transformations closer to the centre of the model ($i = $ $. n-2.n-1.n$) would be mostly devoted to semantic and pragmatic issues like anaphora resolution. inter-sentential dependences. general discourse problems. Again. we do not insist that this old-fashioned segmentation has to be maintained. In fact. transformations F_i's are closely related to the changes in our perception of the universe. These changes are reflected in deriving appropriate representations of involved utterances by the transformations on the Left Side of the Stratified Model. Consequently. we can impose no limits on the quality or quantity of information (linguistic and otherwise) that is to be used by an F_i as long as the desired result is produced. Some

degree of sequentiality of processing will probably be desirable. but the distinctions between different transformations may be defined along some other the traditional dimensions.

A meaning representation language at any level SLA_i will be determined by the transformation F_i leading to it. The better understood a transformation. the more we can say about various details of the meaning representation it builds. Recent work in AI. linguistics. and philosophy touch on what to expect at different SLA_i levels. for example. see Quine (1960). Cresswell (1973). Lewis (1976). etc.. Many of them are listed in the bibliography section at the end of this thesis. We shall not investigate. into these problems here. Instead. we show how the two major theories of language discussed in the last chapter can be accommodated into this framework. It would be interesting also to provide some hints on how to construct the ultimate SLA_n level. First. however. let us discuss briefly the problem of ambiguity which inherently arises for any stratified model. no matter how we project its structure. The problem is present in both the Left and Right Sides of the model.

We assume that at any level SLA_i all expressions of the meaning representation language of this level are unambiguous in terms of the universe UA_i corresponding to SLA_i in the Stratified Model. In other words. if $\alpha_i \in SLA_i$ then there is at most one $\beta_i \in UA_i$ such that $M_i(\alpha_i)=\beta_i$. where M_i is the semantic mapping available at the stratum $<SLA_i, UA_i>$ This condition extends on every stratum for $0 \leqslant i \leqslant n$. The English sentence (or an utterance of it) *Every man loves a woman* is unambiguous in the real world. although it describes more than one different "situation". These "situations" are what the sentence gets unambiguously mapped onto. However. when we try to represent the sentence in first order logic (at some level SLA_i) we obtain at least two different translations each corresponding to a different "situation" at some UA_i. Ambiguity is therefore introduced by transformations. and we should be aware of this.

Let α_i be an expression at level SLA_i. and let $F_{i+1}(\alpha_i)$ be its translation into level SLA_{i+1}. In general $F_{i+1}(\alpha_i)$ will be a set of expressions from SLA_{i+1} that account for α_i ambiguity relative to the characteristic of this level. α_i may be very well ambiguous in some other aspects but it will go

unobserved here. Imagine that for the transformation F_{i+1} there is the inverse transformation F_{i+1}^{-1} such that it returns the set $F_{i+1}(\alpha_1)$ back into α_i at the level SLA_i. We have

$$(2) \quad F_{i+1}(\alpha_i) = \left\{ \alpha_{i+1} \mid F_{i+1}^{-1}(\alpha_{i+1}) = \alpha_i \right\}$$

This explanation is slightly unrealistic as some of the α_{i+1} s will probably be eliminated as impossible (excluded by the SLA_{i+1} s language), or even unlikely (if we admit some probabilistic, or common-sense measures). This situation happens because of a lack of syntactic, semantic, or pragmatic well-formedness. For example, a lexically ambiguous sentence may prove syntactically uniform, etc. So we have to add the obvious constraint to $F_{i+1}(\alpha_i)$ to the effect that, for any $\alpha_i \in SLA_i$,

$$(3) \quad F_{i+1}(\alpha_i) = \left\{ \alpha_{i+1} \mid F_{i+1}^{-1}(\alpha_{i+1}) = \alpha_i \; \& \; \alpha_{i+1} \in SLA_{i+1} \right\}$$

When it comes to the next transformation, F_{i+2}, all α_{i+1} s will be translated by themselves, but we will probably want to keep track of the translation structure built by previous transformations.

More or less the same happens on the Right Side, when approximating the universe. The fact that we discover an ambiguity in a language utterance is a consequence of the observation that the "situation" the utterance refers to in the corresponding universe model is not uniformly perceived, i.e., it is ambiguous. A new universe model is being derived by some G_i in which the old situation is replaced by two or more new distinct situations, although the former situation may still be observed. The corresponding transformation F_i on the Left Side derives appropriate linguistic representations of the ambiguous utterance. As a consequence, we regard a transformation F_i as a linguistic process on an utterance that reflects a new perception of the universe: new structuring, new fragmentation, new generalization, etc. In this sense most of the meaning-preserving string manipulations on a linguistic expression will not be classified as autonomous transformations in the Stratified Model.

Let us now examine the structure of a stratum $< SLA_i, UA_i >$, $0 \leqslant i \leqslant n$, that is, a pair consisting of a meaning representation level and the corresponding universe model within the Stratified Model.

---

† We place the word "situation" into apostrophes here to avoid a reference to the theory of situations.

We have assumed earlier that the language of the level SLA_i, the i-th degree transformation of the source language SL, communicates about UA_i, the i-th degree representation of the original universe U. In other words, there exists a mapping M_i, call this mapping semantics if you like, such that $M_i(SLA_i) = UA_i$. Except for the ultimate mapping $M_n = M$ in the Stratified Model, the relation M_i is too complex to be computed directly. It is quite common in current AI practice that an M_i is attempted for implementation before the ultimate stratum is reached, that is, for $i < n$. In effect we normally describe some submapping $M_i' \subset M_i$ such that

(4) $\quad SLA_i \supset SLA_i' \rightarrow_{M_i'} UA_i' \subset UA_i$

with the problems not covered by M_i' given, at best, some ad hoc solutions. Though undesirable, the method resulted in quite a number of satisfactory, but of limited utility, AI systems, e.g., LUNAR (Woods et al. 1972), PLANES (Waltz et al. 1976), MARGIE (Schank 1975), etc. Our ultimate goal can be roughly verbalized now: to find that ultimate stratum $< SLA_n, UA_n >$ so that M_n could be entirely computed. A hint on what we might expect to find at this stratum may be provided by what we feel should be the ultimate meaning representation level for the source language.

It would be most desirable by the time we reach the level SLA_n when translating some expression α from $SL = SLA_0$ † that we get a single (not a set of possibilities) representation α_n of α at SLA_n if one can be produced at all. That is

(5) $\quad F_n O F_{n-1} O \cdots O F_1(\alpha) = \alpha_n \in SLA_n$

where $F_i O F_j$ is a transformation that combines effects of constituent transformations. The level SLA_n would be therefore the one at which every expression of SL gets a single, unambiguous representation. We must remember that the notion of ambiguity is relative to both the SLA_n's language and UA_n structure. At one possible SLA_n we may get a set A_n of α_n's as the translation of α, but at another, the same set will be considered as an atomic expression. This consideration introduces an element of subjectivity into the translation process which cannot be entirely avoided. In a sense,

---

† Remember that α does not have to be any commonly understood syntactic unit, and may stand for a entire discourse.

therefore, the original problem of finding the mapping between SL and U satisfied (5). When taken at 'face value. language is unambiguous but an utterance can rarely be taken at face value When we utter anything we certainly mean something by that utterance. and if the hearer expects so. as he most probably does. our utterance is potentially ambiguous for him. That means that both speaker and addressee actually compute some steps in the Stratified Model to get the intended and presumed meaning of the utterance respectively. by reaching the same ultimate stratum $<SLA_n, UA_n>$. This condition connotes that not all stages in the Stratified Model are of equal visibility. and some may be passed unnoticed.

3.2. Discussion

Let us examine one possible way to accommodate Montague's PTQ. and Situation Theory in our framework. We show some deficiencies of these theories and suggest a way to correct them. Then we define the extent to which this thesis contributes to the Stratified Model

We take Montague's PTQ as a particular example of a natural language system based on possible world semantics. When compared to the Stratified Model. PTQ exposes a striking imbalance in the degree of development of its Left and Right sides. While the language transformation problems (at least from some level) are well worked out (especially when you add some extensions to the original theory as suggested by Partee (1976). Thomason (1976). Dowty (1976 1981) etc). modelling the universe is virtually neglected. The most natural interpretation is to consider the PTQ's system of possible worlds. as the (alleged) ultimate representation of the universe UA_m. Yet. as we have already pointed out. no intelligible decoding process is suggested to relate possible worlds to reality. nor how to build a proper system of possible worlds in a general case. On the language side we can clearly differentiate three right-most levels. call them SLA_{m-2}. SLA_{m-1}. and SLA_m. with SLA_m being mapped onto UA_m. One can easily identify them with "the fragment". its image in the categorial grammar CAT and IL. respectively. The earlier stages SLA_i for $i < m-2$ are left implicit. PTQ specifies then. except

the mapping M_n between SLA_m and UA_m (IL semantics) two major transformations on the Left Side

The transformation F_{n-1} takes "the fragment" onto the language of p-markers in the categorial gram-

mar Montague (1974f) The translation into IL accounts for the next transformation F_m Figure 3.4

below places PTQ in the Stratified Model scheme

Perhaps the reader has observed that we used some subscript m instead of n $(m < n)$ in denoting

PTQ s final stratum $< SLA_m, UA_m >$ This is because we think that Montague closed the Left Side

too early Montague s system leaves unresolved the problems of inter-sentential dependences and

discourse analysis We suggest some solutions to these problems in this thesis

Let us now turn to the theory of situations From the perspective of our Stratified Model, the

only thing that is clearly visible is the ultimate stratum $< SLA_n, UA_n >$ with SLA_n being the language
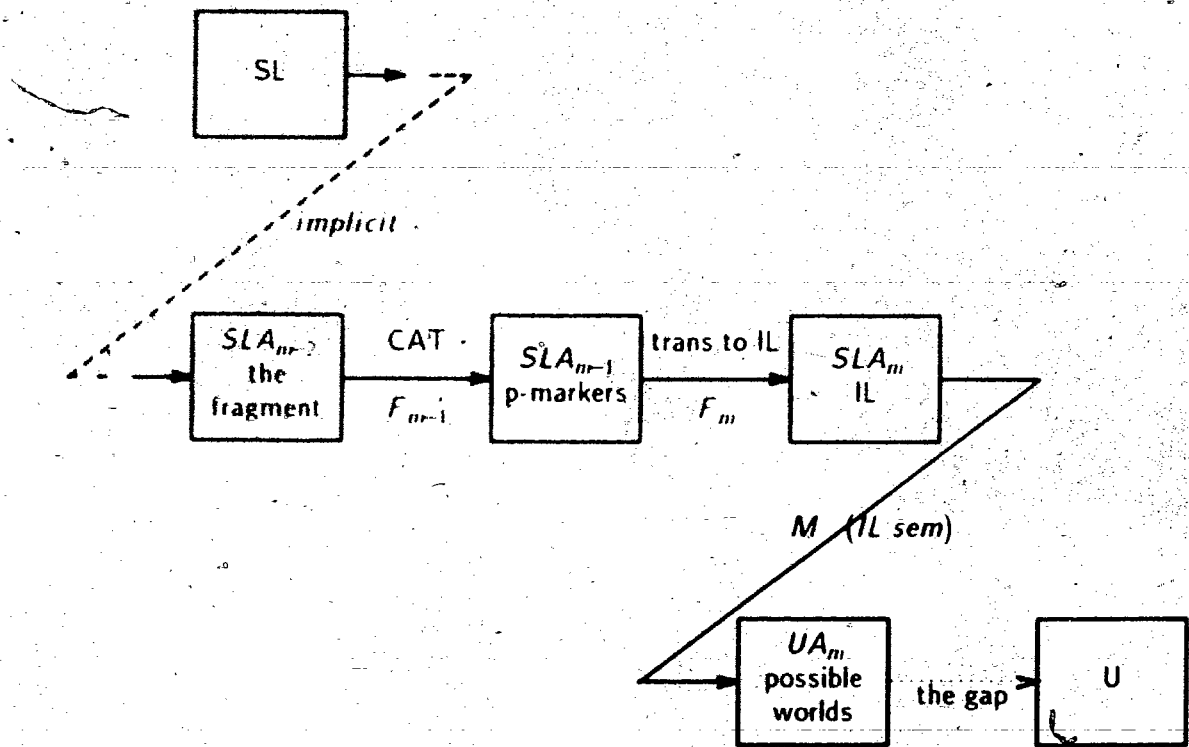
Figure 3.4. PTQ in the Stratified Model

of abstract situations. and UA_n being a system of "real" situations. Barwise & Perry (1983) Although

no specific decoding function is proposed. the structure of real situations as presented by Barwise and

Perry (1983). is rich enough to give one a pretty good idea how to build the Right Side and reach the

universe This provides some insight on how the appropriate transformations G_n, G_{n-1} G_1 may

look. although we won't bother to investigate these problems here On the other hand. the Left Side is

completely unaccounted for. We might pick up some intuition on what to expect in the F_n or perhaps

even F_{n-1} transformation. but this certainly does not close the gap between the source natural

language and the language of abstract situations Finally. let us observe that the ultimate stratum

$<SLA_n$, $UA_n>$ of Situation Theory is very likely to be the right one but as we remember from the

last chapter. the theory still needs to establish formal foundations before one can fully appreciate its

insights Figure 3.5 shows how the Situation Semantics theory looks from the Stratified Model

viewpoint.

Figures 3.4 and 3.5 actually show that both language theories are mutually complementary. even if we

feel that some more SLA_i, $i > m$. levels are required in the PTQ scheme Suppose we can create such a

level SLA_n (after some necessary modifications to the other elements in the model) If we could
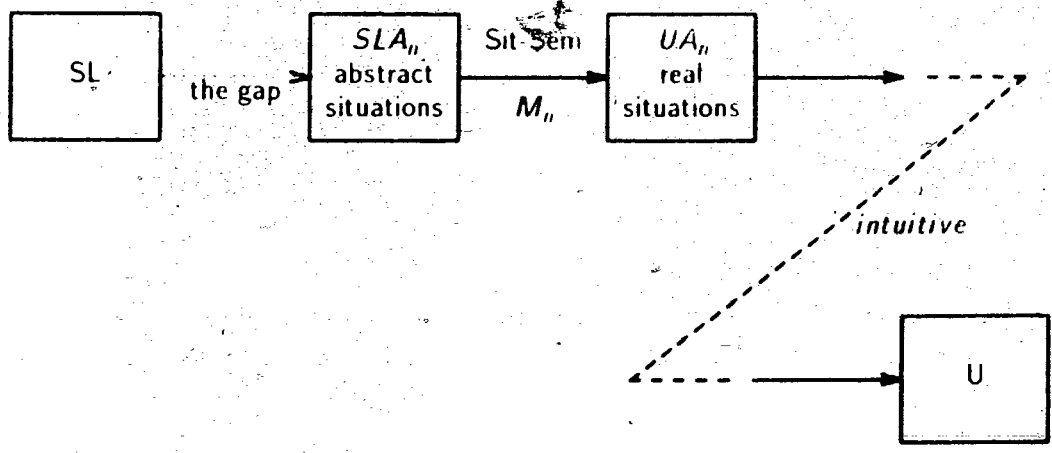
Figure 3.5. Situation Semantics in the Stratified Model

somehow identify the modified PTQ s SLA_n with that of Situation Theory. we would be very close indeed to the first complete Stratified Model! What we need to do is the following. Remove the existing SLA_m level (IL) in the PTQ scheme and replace it by two new levels SLA_m and SLA_{m+1} to account for inter-sentential dependencies and discourse analysis problems, respectively. Then rename levels so that the ultimate stratum becomes $<SLA_n, UA_n>$. These operations may, in fact, involve some changes to some earlier levels SLA_j, $j < m$. in particular to the sentence representations produced by the fragment s categorial grammar. This aspect of the problem will also be discussed in this thesis. Then comes the most difficult step to reach an agreement between SLA_n levels of the modified PTQ and Situation Theory. This should not look so hopeless considering our discussion from the last chapter. Finally, we build the Right Side and adjust both sides to obtain a Stratified Model.

Finally, we come to the last problem we want to mention in this chapter. What we still require for our scheme to work is some new mapping M, the semantics at the $<SLA_n, UA_n>$ stratum. Two possible candidates are in sight. The possible world semantics. and the Situation Semantics. The

from SL
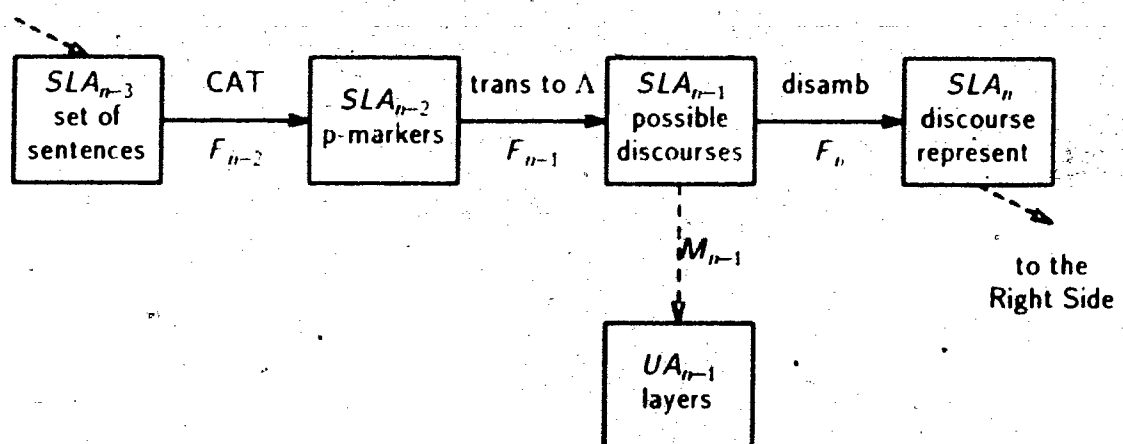
Figure 3.6. The fragment of the Stratified Model built in this thesis.

former looks computationally problematic and of course it cannot make a direct use of the theory of situations. The latter is more promising though it lacks necessary formal foundations. To use the situation theory in an AI system we have to formalize transformations G, and find proper foundations for M. This seems to be a big job indeed. Then maybe we can find another alternative? We think that this thesis takes us towards what we believe is such an alternative

A lot of work remains to be done. We do not provide answers for all of these questions. We suggest two extra levels to be added to a scheme resembling that of PTQ with strong emphasis on the fact that the ultimate meaning representation should be identifiable (at least theoretically) with that of abstract situations. Some elements of the two levels SLA_{n-1} and SLA_n are described in the next three chapters which make the kernel of the thesis. The appropriate transformations F_{n-1} and F_n will be the translating into some λ-categorial language Λ (Chapters 4 and 5) and disambiguating in discourse (Chapter 6). We discuss whether the so defined level SLA_n can be a right choice for the ultimate stratum. We shall also investigate selected problems of universe modelling, especially at the level UA_{n-1}, in connection with our Theory of Names and Descriptions presented in Chapter 5. Figure 3.6 shows the levels of the Stratified Model we discuss in this thesis

Chapter 4

Inter-sentential Dependencies

4.1. Using a λ-categorial Language for Meaning Representation

We assume that we have an arbitrarily selected English subset FMT $\subset$ SL. Let L be that part of the language at level SLA_{n-2} in the Stratified Model where $F_1 \circ \cdots \circ F_{n-2}(FMT) = L$. We concentrate on the translation of some selected example expressions, sentences and paragraphs of L into a typed λ-calculus-based λ-categorial language Λ defined at level SLA_{n-1}. We assume that the reader is familiar with a categorial grammar of a simple fragment of English, such as FMT (for example, the PTQ's fragment discussed in Chapter 2). A number of examples are presented in this and subsequent chapters which are formulated in FMT syntax. The reader is reminded, however, that the language L may differ considerably from FMT, since L is the product of $n-2$ transformations already performed over FMT. Although the transformations from F_1 to F_{n-3} are neglected in this presentation, the transformation F_{n-2}, identified here with the categorial grammar CAT, provides appropriately "parsed" expressions, sentences and paragraphs of FMT. It is not necessary that F_{n-2} be a categorial grammar; perhaps some other syntactic system would be more suitable in practice. Nonetheless the simplicity and elegance of CAT make this grammar most suitable for this presentation. The transformation F_{n-1} which we construct operates on the level SLA_{n-2} to which L belongs.

Some basic concepts are introduced which lay the foundations for further discussion. We sketch the definition of the language Λ as required for this presentation. Λ possesses adequate expressive power to represent discourse meaning at the level SLA_{n-1}. We also speculate that the language promotes computational efficiency. We formulate a number of rules in this work. Many of them are

101

verbalized in Λ. Later when we present our Theory of Names and Descriptions. we shall enrich Λ to cover the multi-layered model of language. and try to establish a correspondence between the two versions.

Definition 1 (Lexicon)

The well-formed expressions of Λ are built of symbols which fall into the following six classes

(a)   VAR of variables: $t.$ $u.$ $x.$ $y.$ z (individual variables). $P.$ $Q.$ $R.$. $C_1.$ C_2.   (predicate variables):

(b)· CON of constants: $a.$ $b.$ $c.$. $A.$ $B.$ D.

(c)   PAR of parentheses: $(.)$

(d)   LAM of lambda-abstractor λ.

(e)   LOP of logical operators: $\&.$ $\cup.$ $\supset.$ $\equiv.$ $\neg.$

(f)   QUA of quantifier symbols $\exists.$ $\forall.$

The nature of the sets VAR and CON is not uniform In fact a structure of types has been superimposed over the language so that every element of VAR and CON is assigned to some type   Let α be such a type. then by VAR_α and CON_α it is understood the set of variables of type α and the set of constants of type α respectively.

Definition 2 (Types)

The set of types of expressions of Λ is the smallest set TYPES such that

(a) $t.$ $e \in$ TYPES:

(b)   for any $\delta.$ $\sigma \in$ TYPES. $\sigma/\delta \in$ TYPES. $\square$

Here t and e are the basic types which are not in the form σ/δ for any $\sigma.$ $\delta \in$ TYPES. With this concept of type we can define the notion of *well-formedness* of expressions in Λ

Definition 3 (Syntax)

Let $\alpha, \beta \in$ TYPES. The set of well-formed expressions of type α is the smallest set WFE_o such that

(a) if $x \in VAR_o$ then $x \in WFE_o$;

(b) if $a \in CON_o$ then $a \in WFE_o$;

(c) if $E \in WFE_o$ and $x \in VAR_\beta$ then $\lambda x E \in WFE_{\alpha/\beta}$;

(d) if $E_1 \in WFE_{\alpha/\beta}$. $E_2 \in WFE_\beta$ then $(E_1 \, E_2) \in WFE_o$;

(e) if E_1. $E_2 \in WFE_o$ then $(E_1 = E_2) \in WFE_t$.

(f) if E_1. $E_2 \in WFE_t$. $u \in VAR$ then $\neg E_1$. $(E_1 \, \& \, E_2)$. $(E_1 \cup E_2)$. $(E_1 \supset E_2)$. $(E_1 \equiv E_2)$.

$\forall u E_1$. $\exists u E_1 \in WFE_t$ $\square$

Let L_{n-2} and L_{n-1} be fully developed languages defined at levels SLA_{n-2} and SLA_{n-1} respectively. Our first effort is to describe the transformation F_{n-1} such that $F_{n-1}(L_{n-2}) = L_{n-1}$. or more precisely. to formulate a collection of rules $R_{n-1} = \left\{ R_{(n-1)_1}, \ R_{(n-1)_2} \qquad \right\}$. $R_{n-1} \subseteq F_{n-1}$ such that $R_{n-1}(L) = \Lambda$. By saying that the collection of rules R_{n-1} is a subset of the transformation F_{n-1} it is understood that the sum of domains of rules within this set does not exhaust the domain of the transformation. Note that the transformation F_{n-1} and any other transformation F_i. $1 \leq i \leq n$ is not a function in the set theoretic sense. In general. an application of a transformation to a language expression may result in more that one different translations. The first rule we present here summarizes most of the translation rules suggested by Montague (1974f). Partee (1976). and Dowty (1976) with Λ used in place of Intensional Logic (IL). We assume the straightforward correspondence between categories of L and types of Λ. i.e.. if α is a basic category in L. then $F_{n-1}(\alpha)=\alpha$ is a type in Λ: if α is a derived category: σ/δ. $\sigma//\delta$. ... then $F_{n-1}(\alpha)=\sigma/\delta$ is a type in Λ. The rule is stated below. The reader is remainded that $B(\alpha)$ and $E(\alpha)$ stand for the set of basic expressions of category α within L and the set of all expressions of category α within L. respectively.

RULE 1 (The Basic Translation Rule)

(i) Let α be any category in L, different than $T = t/(t/e)$. If $a \in B(\alpha)$ then $F_{n-1}(a) \in CON_{F_{n-1}(\alpha)}$.

(ii) If $a \in B(T)$ then $F_{n-1}(a) = (\lambda P \ (P \ a'))$, where $a' \in CON_e$ and $P \in VAR_{t/e}$.

(iii) For any categories α, β of L, if σ is any of the following: $\alpha/\beta, \alpha//\beta$, then if $E_1 \in E(\sigma)$ and $E_2 \in E(\beta)$ then $F_{n-1}(<E_1, E_2>) = E_1 E_2 \in WFE_{F_{n-1}(\alpha)}$, where $< \cdots >$ is the syntactic operation of L.

(iv) If $E \in E(t//e)$ then

$$F_{n-1}(a/an \ E) = (\lambda Q \ (\exists x \ ((F_{n-1}(E)) \ x) \ \& \ (Q \ x)))$$

$$F_{n-1}(every \ E) = (\lambda Q \ (\forall x \ ((F_{n-1}(E)) \ x) \supset (Q \ x)))$$

$$F_{n-1}(the \ E) =$$

$$(\lambda Q \ (\exists x \ ((F_{n-1}(E)) \ x) \ \& \ (C \ x) \ \& \ (\forall y \ (((F_{n-1}(E)) \ y) \ \& \ (C \ y)) \supset (x=y)) \ \& \ (Q \ x)))$$

where $x, y \in VAR_e$, $Q \in VAR_{t/e}$, and C is a context, $C \in WFE_{t/e}$. $\Box$

Example 1

Suppose we want to translate the L sentence

(1) *Bill interviewed every applicant*

into Λ. Suppose also the following translations hold †

(i) *interview* $\rightarrow i$

(ii) *applicant* $\rightarrow a$

(iii) *Bill* $\rightarrow (\lambda P \ (P \ B))$

The constants $i, a,$ and B above belong to types $TV=(t/e)/(t/(t/e))$, t/e, and e, respectively. Applying steps (iii) and (iv) of Rule 1 we obtain (based on a correct syntactic analysis in CAT)

(iv) $1 \rightarrow (\forall x \ (a \ x) \supset (i \ B \ x)) \Box$

---

† We shall use the symbol $\rightarrow$ as the translation operation symbol in place of Rule 1's F_{n-1}.

Example 2

Suppose we have the following translations.

(i) *man* $\rightarrow$ *m*

(ii) *woman* $\rightarrow$ *w*

(iii) *loves* $\rightarrow$ *l*

Here, again, the constants m, w, and l are respectively of types t/e, t/e, and TV. The two possible grammatical analyses of the sentence

(2) *Every man loves a woman*

lead to two different translations called the *weak* and *strong* readings of (2) respectively. These are

(iv) $2 \rightarrow (\forall x \, (m \, x) \supset (\exists y \, (w \, y) \, \& \, (l \, x \, y)))$

and,

(v) $2 \rightarrow (\exists y \, (w \, y) \, \& \, (\forall x \, (m \, x) \supset (l \, x \, y))). \quad \Box$

We now focus on selected examples of two-sentence "stories" and try to discover and formalize referential interdependencies between them and the conditions in which such dependencies arise. Restricting the discussion to two sentence "stories" avoids, at this stage, most of the problems of *where to look for the reference*, thus we concentrate entirely on the question *how to get the reference*. A good example of the former was given by Nick Cercone (personal communication) in the form of the three-sentence dialogue (in a restaurant, to a waiter):

speaker-A: *I'll have pepsi.*

speaker-B: *I'll have nothing.*

speaker-C: *I'll have the same.*

Although we are aware of this kind of problem, for now we consider only situations where a reference, if can be made at all, has a unique antecedent. The fact that we restrict our discussion to two-sentential paragraphs should not be taken literally, especially when interpreting the translation rules

(from Rule 2 onward). Most rules will be written using two hypothetical "sentences" S_1 and S_2 where the latter contains a reference element to an object mentioned in the former. The sentence S_1 will normally be considered to establish a context for making that reference. In fact it is not necessary that any single such sentence exists. We assume only that at the time S_2 is expressed there is enough context information accumulated by different means so that the reference can properly be made. i.e., the referent is uniquely identified. In our idealized model all this is reduced to the situations described by the sentence S_1 or its elements. but in general, S_1 need neither be a single sentence nor a collection of sentences, and may contain information acquired from other sources (observation. knowledge base. beliefs, etc).

4.2. The THE Determiner

The traditional translation for *the* determiner which can be found in numerous works, see for example Montague (1974f). Lewis (1976). Partee (1976). tends to be

If a is a common noun which translates to a'. and $F(a)$ denotes *the a*. then $F(a)$ translates into
$\hat{P}[\exists y \ (\forall x \ [a'(x)\equiv x=y]) \ \& \ P(y))]$

When mapped into our λ-notation the formula above becomes:

- *the a* $\rightarrow$ $(\lambda P \ (\exists y \ (\forall x \cdot (a \ x) \equiv (x=y)) \ \& \ P(y)))$

As explained by Partee (1976).

[*the king* denotes] the set of all properties such that there is a unique entity which is a king. and he has those properties.

According to the type structure of L we can obtain the translation of the *the* determiner from the above as:

(THE1)

the $\rightarrow$ $(\lambda Q \ (\lambda P \ (\exists y \ (\forall x \ (Q \ x) \equiv (x=y)) \ \& \ (P \ y))))$

This translation has two obvious weaknesses:

(P1) It presupposes the existence of the object it determines;

(P2) It assumes we can always find an environment in which the entity is unique (i.e., unique king, man, queen, etc.).

We show that translation THE1 fails when a wider than single sentence context is used.

Example 1

Let us consider the following "story".

(1a) *John interviewed a man.*

(1b) *The man killed him.* (i.e., *John,* we shall assume this henceforth)

According to the traditional translation of *the* by the formula THE1, (1b) translates into

- $(1b) \rightarrow (\ ]y\ (\forall x\ (m\ x) \equiv (x=y))\ \&\ (k\ y\ J)).$

  where *killed* $\rightarrow$ *k, man* $\rightarrow$ *m.*

This translation is questionable. Notice that it is not the fact of being a man which makes the entity unique. Observe also that even a context-less sentence

- *The queen is wealthy*

cannot be properly understood without a clear reference to the queen in question. i.e.,

- *The queen such that P(the queen) is wealthy.*

where P belongs to $\{I\ know\ something\ of\ her,\ I\ can\ see\ her,\ ...\}$. †

Thus it appears that the use of a definite description requires some **context setting situation** C where its reference could be validated. The appropriate *the* translation could be modified to

---

† There are sentences closely resembling the above which do not require restrictive context-setting situations because the use of the *the P* phrase is generic in them as in: *The tiger lives in the jungle.* Refer. however. to Vendler (1971) for further discussion. In Chapter 5 when we attempt to construct a general theory of names and descriptions we will see where to place the so-called "generic" and other attributive uses of nominal phrases.

(THE2)

> **the** $\rightarrow$ $(\lambda P (\lambda C (\lambda Q (\exists x (P\ x) \ \& \ (\forall y ((P\ y) \ \& \ (C\ y)) \equiv (x{=}y)) \ \& \ (Q\ x)))))$

Using THE2, the translation of sentences (1a) and (1b) become

- 1a $\rightarrow$ $(\exists x (m\ x) \ \& \ (i\ J\ x))$

- 1b $\rightarrow$ $(\exists x (m\ x) \ \& \ (\forall y ((m\ y) \ \& \ (i\ J\ y)) \equiv (x{=}y) \ \& \ (k\ x\ J))$

At first glance the formula THE2 seems overstated. One can observe that a definite description may be used in a more or less "definite" manner so that, at one extreme, we do not need any external context to refer unambiguously to an object. Subsequently the description

- The man in a grey coat I saw yesterday in the library

still seems amenable to proper translation by the formula THE1. Let us use, following Barwise and Perry (1983), the term **described situation** for the situation referred to by the utterance itself We see that in the example above the described situation is contingently the same as the context setting situation. Therefore we can consider formula THE1 as just a special case of the more general formula THE2. Consider, however another controversial example

<u>Example 2</u>

Suppose we have three different "stories" created by pairs of sentences (2a. 2b). (2a. 2c). and (2a. 2d).

(2a) *John interviewed a man*

(2b) *The bastard killed him*

(2c) *The employee killed him.*

(2d) *The woman killed him.*

(We assume that *him* refers to *John*, even if it needs not be the case in general.) Even a cusory glance at these sentences should convince us that the facts $(P\ x)$ and $(C\ y)$ in the THE2 translation formula

cannot be specified arbitrarily, and actually their mutual compatibility has a significant influence on the anaphora referent-taking process. The *the*-referenced object must belong to a class at least as broad as the the intended referent in the context. So in (2b) *the bastard* can be identified with *a man* while the same process for *the employee* in (2c) is less obvious. In (2d) it is apparent that *the woman* and *a man* from (2a) operate in distinct contexts. This is not to say that *the woman* of (2d) cannot be taken as referring to the man described in (2a). If that was the case, the fragment would assert of some individual that it is both a man and a woman. Depending on a particular interpretation of these predicates. and certain other circumstances which we discuss in Chapter 5. such a reading may be discarded as inconsistent. These considerations lead to the final *the* translation THE3.

(THE3)

the $\rightarrow$ $(\lambda P\, (\lambda C\, (\lambda Q\, (\exists x\, (P\, x)\, \&\, (C\, x)\, \&\, (\forall y\, ((P\, y)\, \&\, (C\, y))\, \supset\, (x=y))\, \&\, (Q\, x)))))$

Here the facts $(P\, x)$. $(Q\, x)$. and $(C\, x)$ give characteristics of the referenced object. The part under the universal quantifier emphasizes the uniqueness of the x under the context C. The biconditional in THE2 can be dropped now, as we explicitly assert P and C of x. Observe that the literal $(P\, y)$ in the part under the universal quantifier is often insignificant for fixing a unique reference for the object. Examine, for example, the story consisting of (2a) and (2b). The same cannot be said, however, of the instance $(P\, x)$ outside the scope of $\forall$. In this case we acquire some additional knowledge about the already selected individual. In the rest of this discussion we shall often drop the literal $(P\, y)$ whenever it does not lead to an ambiguous situation.

The next question to ask is how can we uniformly establish the context of C? The following two cases are clearly perceptible.

(S1) The context is unknown as in *The queen (I can see her) is very wealthy*. We must employ pragmatics to resolve this.

(S2) The context is known from a previous statement with *a* determiner. Having

$(\exists x\, (P\, x)\, \&\, (Q\, x))$ as the context setting sentence, we get the context for x as

$$(\lambda x \ (P \ x) \ \& \ (Q \ x))$$

We are ready now to present the first formal rule. the **Perfect Context Translation Rule** for translating two sentence paragraphs.

RULE 2 (Perfect Context Translation Rule)

An object u referenced by *the* has been mentioned previously in a *de re* context. thus its existence is presupposed. Let $S_1(u)$ be the context sentence which mentions u. Let $S_2(u)$ be the sentence in question. We have

(i) $S_1(a \ P) \ \rightarrow \ (\exists u \ (P \ u) \ \& \ (F \ u))$

(ii) $S_2(\text{the} \ P_1) \ \rightarrow \ (\exists u \ (C \ u) \ \& \ (\forall x \ ((P_1 \ x) \ \& \ (C \ x)) \supset (x=u)) \ \& \ (P_1 \ u) \ \& \ (F_1 \ u))$

The context C is derived from S_1 as $(\lambda u \ (P \ u) \ \& \ (F \ u)) \ \square$

Two examples which employ the THE3 *the* translation illustrate Rule 2 The first is simply a repetition of Example 1. from the beginning of this section. in different environment The other example shows the power of the THE3 formula when dealing with multiple link environment

Example 1 (modified)

Let (1a) be as given before. i e..

• $1a \ \rightarrow \ (\exists x \ (m \ x) \ \& \ (i \ J \ x))$

We derive the context C from (1a) as

• $(\lambda x \ (m \ x) \ \& \ (i \ J \ x))$

The incremental translation of (1b) follows.

•   *the man* $\rightarrow$

     $(\lambda C \ (\lambda Q \ (\exists x \ (m \ x) \ \& \ (C \ x) \ \& \ (\forall y \ ((m \ y) \ \& \ (C \ y)) \supset (x=y)) \ \& \ (Q \ x))))$

     the man (*whom John interviewed*) $\rightarrow$

$$(\lambda Q\ (\ |x\ (m\ x)\ \&\ ((\lambda x_1\ (m\ x_1)\ \&\ (i\ J\ x_1))\ x)\ \&$$

$$(\forall y\ ((m\ y)\ \&\ ((\lambda x_1\ (m\ x_1)\ \&\ (i\ J\ x_1))\ y))\ \supset\ (x{=}y))\ \&\ (Q\ x)))\ \rightarrow$$

$$(\lambda Q\ (\ |x\ (m\ x)\ \&\ (i\ J\ x)\ \&\ (\forall y\ ((m\ y)\ \&\ (i\ J\ y))\ \supset\ (x{=}y))\ \&\ (Q\ x)))$$

the man (whom   ) killed John $\rightarrow$

$$(\ |x\ (m\ x)\ \&\ (i\ J\ x)\ \&\ (\forall y\ ((m\ y)\ \&\ (i\ J\ y))\ \supset\ (x{=}y))\ \&\ (k\ x\ J))$$

where *killed John* $\rightarrow$ $(\lambda v\ (k\ v\ J))$ as before. □

Observe that if literals P and P_1 in the formula THE3 are identical, as will often happen in practice, we can drop P_1 to make the translation shorter. We use this convention in the following example. Note, however, that if *the woman* was used in place of *the man* in (1b) the latter sentence would translate as (if we insisted that *the woman* and *the man* co-referred):

- (1b) $\rightarrow$ ($|x\ (m\ x)\ \&\ (w\ x)\ \&\ (i\ J\ x)\ \&\ (k\ x\ J)\ \&$

$$(\forall y\ ((m\ y)\ \&\ (w\ y)\ \&\ (i\ J\ y))\ \supset\ (x{=}y)))$$

This translation can be discarded right away when we add an interpretation to our language that precludes an individual from being both a man and a woman at the same time. In a practical implementation this feature will constitute an important criterion for selecting proper referents. It must be noted, however, that there are circumstances in which being a man and a woman is not contradictory, even without a change to the interpretation. We discuss such situations in Chapter 5.

Let us now examine how the formula THE3 can be used to translate a variation of so-called "donkey sentences" which are widely discussed by Heim (1982) and Hornstein (1984).

Example 3

We show a detailed translation of (3b) in the context of (3a) below. It is assumed here that *the man* and *the donkey* in (3b) are used to refer to the man (the donkey owner) and the donkey as addressed in (3a), respectively. Observe that this presupposition, though most feasible, is not the only way to interpret this discourse. See also section 4.3 for more discussion.

(3a) *A man owns a donkey*

(3b) *The man beats the donkey.*

This time the context must be drawn for both *the man* and *the donkey* from the same source. i.e. from (3a). but from the different points of view. It is interesting to see how the translation "merges" the contexts in one common reference. Let us start by establishing partial translations for sentence fragments.

- *the man* →

  $*(\lambda Q\ (\exists x\ (m\ x)\ \&\ (C\ x)\ \&\ (\forall y\ (C\ y)\ \supset\ (x=y))\ \&\ (Q\ x)))$

  the donkey →

  $(\lambda Q\ (\exists v\ (d\ v)\ \&\ (C_1\ v)\ \&\ (\forall y\ (C_1\ y)\ \supset\ (v=y))\ \&\ (Q\ v)))$

  where *donkey* → *d*

Assume *beat* → *b*. then

- *to beat the donkey* →

  $(\lambda s\ (\exists v\ (d\ v)\ \&\ (C_1\ v)\ \&\ (\forall y\ (C_1\ y)\ \supset\ (v=y))\ \&\ (b\ s\ v)))$

Now *the man* is supplied as the subject to *beat*.

- *the man beats the donkey* →

  $(\lambda C\ (\lambda C_1\ (\exists x\ (\exists v\ (m\ x)\ \&\ (C\ x)\ \&\ (\forall y\ (C\ y)\ \supset\ (x=y))\ \&\ (d\ v)\ \&\ (C_1\ v)\ \&$

  $(\forall y\ (C_1\ y)\ \supset\ (v=y))\ \&\ (b\ x\ v)))))$

In this point we cannot proceed further without establishing the contexts C and C_1. The translation of (3a) is:

- $3a\ \rightarrow\ (\exists x\ (\exists v\ (m\ x)\ \&\ (d\ v)\ \&\ (own\ x\ v)))$

According to Rule 2. both contexts are derived as:

- $C\ =\ (\lambda x\ (\exists v\ (m\ x)\ \&\ (d\ v)\ \&\ (own\ x\ v)))$

i e. from the man point of view: *a man who owns a donkey* and

- $C_1 = (\lambda v \; (\exists x \; (m \; x) \; \& \; (d \; v) \; \& \; (own \; x \; v)))$

i.e. from the donkey perspective: *a donkey which is owned by a man.* Supplying C and C_1 into the translation of (3b) we have got the following

- 3b →

  $(\exists x \; (\exists v \; (m \; x) \; \& \; (\exists v_1 \; (m \; x) \; \& \; (d \; v_1) \; \& \; (own \; x \; v_1)) \; \&$

  $(\forall y \; (\exists v_1 \; ((m \; y) \; \& \; (d \; v_1) \; \& \; (own \; y \; v_1)) \supset (x=y)) \; \&$

  $(d \; v) \; \& \; (\exists x_1 \; (m \; x_1) \; \& \; (d \; v) \; \& \; (own \; x_1 \; v)) \; \&$

  $(\forall y \; (\exists x_1 \; ((m \; x_1) \; \& \; (d \; y) \; \& \; (own \; x_1 \; y)) \supset (v=y)) \; \& \; (b \; x \; v)))$

This formula is complete but it is somehow too complicated. and obviously contains redundant information. Observe that from the subformula $(\exists v_1 \ldots)$. and the subformula $(\forall y \; (\exists x_1 \ldots))$ we can get $v=v_1$. while from subformulae $(\exists x_1 \ldots)$. and $(\forall y \; (\exists v_1 \ldots))$ we have $x=x_1$. Therefore we reduce our result accordingly. obtaining

- 3b →

  $(\exists x \; (\exists v \; (m \; x) \; \& \; (d \; v) \; \& \; (own \; x \; v) \; \&$

  $(\forall y \; (\exists v_1 \; ((m \; y) \; \& \; (d \; v_1) \; \& \; (own \; y \; v_1)) \supset (x=y)) \; \&$

  $(\forall y \; (\exists x_1 \; ((m \; x_1) \; \& \; (d \; y) \; \& \; (own \; x_1 \; y)) \supset (v=y)) \; \& \; (b \; x \; v)))$

That is almost our desired result. We still have two separate context descriptions generated from two different points of view. i.e.. from the perspectives of *the man* and *the donkey* respectively. Taking these different points of view into consideration. further reduction yields

- 3b →

  $(\exists x \; (\exists v \; (m \; x) \; \& \; (d \; v) \; \& \; (own \; x \; v) \; \&$

  $(\forall u \; (\forall y \; ((m \; y) \; \& \; (d \; u) \; \& \; (own \; y \; u)) \supset ((x=y) \; \& \; (v=u)))) \; \& \; (b \; x \; v)))$

This translation can be paraphrased as *The man who owns the donkey beats it.* □

The Perfect Context Translation Rule explains some aspects of translating two sentence "stories". but as we shall see it accounts for only the most straightforward referential situations which do not involve either imperfect or attitude report constructions

4.3. Imperfect Contexts

We distinguish two very general classes of verbs found in natural language sentences. These are **imperfect verbs** like: *seek. want to α. go. build. imagine* and **perfect verbs** like: *find. come. have. have written. have imagined* ... An informal definition of imperfect verbs follows.

<u>Definition</u>

A verb will be called *imperfect* if the immediate effects of the action or state described by this verb last at most as long as the action or state itself does. and its results on the surrounding world cannot be determined before the action or state is committed $\square$

We call such verbs *imperfect* for their ability to create descriptions of situations which may never come to exist When such a situation is used as a context we shall refer to it as an *imperfect context* for the same reason. This ability to produce imperfect contexts is not always a permanent property of a verb and may be limited to some of its forms only (such as future tense. for instance). In such a case our term refers to these selected forms only

The characteristics of imperfect verbs can be summarized as follows

(a) They have no permanent influence on on the situation surrounding an utterance *want to marry. must have.* ...:

(b) They can be committed. however. when turned into a perfect form: *have married. have.* ...

(c) They (but not only they) can create non-referential translations of sentences:

(d) An imperfect verb *v* can be decomposed into an imperfect operator $\tilde{v}$. and the perfect form $\hat{v}$

Thus $seek = try$, $seek = find$. †

(e)   Complement taking imperfect verbs act as *imperfectness* operators on the complement. and its

main verb. creating compound imperfect verbs. A special consideration will be given to the verbs

want and *must*.

The *imperfectness* of imperfect verbs may be further stressed by contrasting them with the attitude

report verbs like *see that*. *imagine that* or *believe that* which can also create non-referential readings

of sentences. Observe that we must use an imperfect verb in (1b). like *must*, to maintain a possible

non-referential reading of this sentence in context of (1a).

(1a) John *wants* to marry a unicorn.

(1b) The unicorn *must* have a pink tail.

This is not. however. the case in (2b) as read in context of (2a).

(2a) John *imagines that* a unicorn lives in the park.

(2b) The unicorn has a pink tail.

This is because *imagines that* as a perfect attitude report verb creates an abstract situation that of

the John's image which survives the utterance of (2a) and can be subsequently referred to directly by

(2b). Imperfect verbs do not possess this property. We discuss attitude report contexts in section 4.4.

Table 4.1 gives some examples of imperfect verbs and their perfect counterparts. Observe how

the context in which an imperfect verb is used influences the *perfecting* operation. Notice that an

imperfect verb can be applied to a complement with another imperfect verb as in

•   *John wants to seek a queen*.

thus raising the *imperfectness* level. Two *perfecting* operations must be performed on the last

example to get the perfect form of *John has found a queen*.

---

† In fact. the particle "to" is not a part of the imperfect operator such as "try to" or "want to", etc. In contrast with Montague's (1974) classification of these verbs in the category (t/e)//(t/e) we need imperfect operators in the category (t/e)/t. with the particle "to" being a part of the complement phrase.

Example 4

Assume the following translations.

●● $marry \rightarrow (\lambda s\ (\lambda x\ (m\ s\ x)))$

●● $want \rightarrow (\lambda s\ (\lambda x\ (\lambda F\ (w\ s\ (F\ x)))))$

It should not be difficult to see that processing the sentence

(4a) *John wants to marry a queen.*

leads to two different translations. We divide the derivation into several steps for greater readability. The first derivation is *de dicto*, the second is *de re*. Observe in what order the λ-reductions take place in both derivations. This ordering is consistent with our intuition.

●●— $[(\lambda P\ (P\ J)).\ (\lambda s\ (w\ s\ (F\ x)))] \rightarrow (w\ J\ (F\ x))))$. which means *John wants him_k to F*

   $[(\lambda P\ (P\ J)).\ (\lambda x\ (w\ J\ (F\ x))))] \rightarrow (w\ J\ (F\ J)))$. which means *John wants John to F*

   $[(\lambda F\ (w\ J\ (F\ J))).\ (\lambda y\ (\exists x\ (q\ x)\ \&\ (m\ y\ x)))] \rightarrow$

   $(w\ J\ ((\lambda y\ (\exists x\ (q\ x)\ \&\ (m\ y\ x)))\ J) \rightarrow$

   $(w\ J\ (\exists x\ (q\ x)\ \&\ (m\ J\ x)))$. which is the desired result

Verb Forms		
imperfect	**perfect form**	**possible imperfect operator**
seek	find	try (to)
go	come	try (to)
go to α	α	go (to)
want to α	α	want (to)
wish to α	α	wish (to)
imagine	have imagined	try (to)
be building	have built	try (to)

Table 4.1. Imperfect verbs and their perfect forms

The *de re* derivation is given below.

- $[(\lambda F (w \ s \ (F \ x))) . (\lambda y \ (m \ z \ y)))] \rightarrow (w \ s \ ((\lambda y \ (m \ z \ y))) \ x)) . \rightarrow$

  $(w \ s \ (m \ z \ x))$. which means he_m wants him_k to marry her_n

  $[(\lambda Q \ (\exists u \ (q \ u) \ \& \ (Q \ u))) . (\lambda x \ (w \ s \ (m \ z \ x))))] \rightarrow$

  $(\exists u \ (q \ u) \ \& \ ((\lambda x \ (w \ s \ (m \ z . x))))) \ u)) \rightarrow$

  $(\exists u \ (q \ u) \ \& \ (w \ s \ (m \ z \ u)))))$. i.e., he_m wants him_k to marry a queen

  Then substituting J for z (as before) we have

  $(\lambda s \ (\exists u \ (q \ u) \ \& \ (w \ s \ (m \ J \ u))))$. i.e., he_m wants John to marry a queen

  and finally, supplying the sentence subject for J, the result is obtained as expected

  $(\exists u \ (q \ u) \ \& \ (w \ J \ (m \ J \ u)))$. $\square$

Let us see now how the phenomenon extends to two sentence paragraphs.

Example 5

Consider the following "story".

(5a) *John wants to marry a queen.*

(5b) *The queen must be wealthy*

As the reader has perhaps already observed the existence of the queen in (5b) is not necessarily presupposed as a consequence of the possible *de dicto* reading of (5a). Notice also that (5b) would have completely different meaning when considered without the context supplied by (5a). We can paraphrase (5a) and (5b) as (in a possible reading)

- *The queen John would eventually marry. if any, must be wealthy*

or

- *John wants to marry a wealthy queen.* $\square$

Examining Example 5 and other similar "stories" we can differentiate two reference situations. The situation where the context setting sentence has its referential reading (i.e., there exists a particular

queen John, wants to marry) is correctly represented by the Perfect Context Translation Rule. This rule, however, cannot be used when both sentences have their non-referential readings. To account for non-referential readings in imperfect contexts we formulate a new rule called **Imperfect Context Translation Rule** given below.

RULE 3 (Imperfect Context Translation Rule)

An object u referenced by *the* has been recently mentioned in a *de dicto* environment, i.e., its existence is not assumed. Let S_1, S_2 be defined as in Rule 2. Then

(i) $S_1(a\ P) \rightarrow (\text{imp}\ (\exists u\ (P\ u)\ \&\ (F\ u)))$ †

where **imp** is the imperfect operator such that it is derived from the imperfect verb of S_1.

(ii) $S_2(\text{the}\ P_1) \rightarrow (\text{imp}_1\ (\exists u\ (C\ u)\ \&\ (\forall x\ ((P_1\ x)\ \&\ (C\ x))\ \supset\ (x{=}u)))\ \&\ (P_1\ u)\ \&\ (F_1\ u)))$

Where **imp$_1$** is the imperfect operator of S_2, and the context C is derived from S_1 as $(\lambda u\ (P\ u)\ \&\ (F\ u))$. □

If a sentence with an imperfect verb has a referential reading in the form

• $(\exists x\ (P\ x)\ \&\ (F'\ x))$

then the non-referential reading featured in Rule 3 is obtained by realising that $\text{imp} = F'$, $F = F'$. To support the formulation of Rule 3 recall sentences (5a) and (5b) from Example 5. Notice that using the imperfect verb *must* (or *will*, with the same effect) in (5b) we extend the "imperfectness" of (5a), and at the same time the *de dicto* reading, on (5b). Notice further that if we used a perfect verb in (5b) as in

(5c) *The queen is wealthy*

---

† Here **imp** is a higher-order operator variable which stands for the imperfect operator (classified into category (t/e)/t) and other sentence elements which are not under the scope of this imperfect operator, if applicable. Thus in non-referential translation of (5a) above "John wants" → **imp**.

we would resolve the *de dicto/de re* ambiguity and **both** sentences would have only their *de re* readings. Thus the presence of an imperfect construction in (5*b*) is essential for preserving (5*a*)'s *de dicto* reading and for extending it over (5*b*). It should also be clear that Rule 3 requires that the object referenced in the current utterance has non-referential status in an imperfect context. In other words, the passage

(5*a*') *John married a queen.*

(5*b*) *The queen must be wealthy.*

has only a referential translation with *must be wealthy* considered as a perfect construction. †

The formal *de dicto* translation of (5*b*) in context of (5*a*) can be sketched as follows.

step1

Select a reference context from (5*a*)'s translation. i.e. from $(w \; J \; (\exists x \; (q \; x) \; \& \; (m \; J \; x)))$ *(de dicto)* We select $(\lambda x \; (q \; x) \; \& \; (m \; J \; x))$.

step2

Base translation of (5*b*) is (according to (ii))

$(\lambda C \; (must \; (\exists u \; (C \; u) \; \& \; (\forall x \; ((q \; x) \; \& \; (C \; x)) \supset (x=u)) \; \& \; (wealthy \; u) \; \& \; (q \; u))))$

step3

Apply the λ-expression obtained in step 2 to that selected in step 1. After simple refinement we get the final translation of (5*b*) as

$5b \; \rightarrow \; (must \; (\exists y \; (q \; y) \; \& \; (m \; J \; y) \; \& \; (wealthy \; y) \; \& \; (\forall x \; ((q \; x) \; \& \; (m \; J \; x)) \supset (x=y))))$

To assess truth conditions of the above formula we would have to determine the semantics of the operator *must* first. This cannot be done, however, until we finally define the corresponding universe model at UA_{n-1} from which formulas like the one above take their interpretations. see Tarski (1935). In Chapter 5 we attempt to build UA_{n-1} but the semantics of *must*, and other second order operators of Λ, are nevertheless left unspecified. The question whether *must* is to be interpreted as the

---

† The reader is remainded that we consider only singular interpretations of nominal descriptions in this chapter.

necessity operator, and thus allowing for inferences like $must(\alpha) \supset \alpha$, is not relevant at this stage. Had it been the case, (5b) would have been reduced to a referential reading allowing for only a referential interpratation of the discourse in Example 5. See also the discussion of Case III in section 4.4.

Nonetheless, Partee (1972) argues that there exists a class of contextual situations closely resembling that of Example 5 where no non-referential reading is possible. According to Partee if the underlined phrases in the following paragraphs are to co-refer they must also have referential interpretations.

(A)   *John wants to marry* a queen.

     Bill wants to marry the queen *too*.

(B)   *John is looking for* a pen.

     Bill is looking for it *too*.

If (A) and (B) were indeed the counter-examples, Rule 3 would have to be rejected or at least reformulated in some narrower sense. But this is not the case. Both examples are somewhat unfortunate because they describe identical attitudes toward an individual which originate from different sources; and this is further stressed by the use of the word *too*. That is why the non-referential readings are so deeply hidden from our intuition. They exist however, and are quite legitimate. If we consequently use Rule 3 then we can interpret the second sentences of (A) and (B) respectively as

- *Bill wants to marry the queen John marries (if any) too (no matter who she is)* [because, say, Bill always tries to do what John is doing]

- *Bill is looking for the pen John (eventually) finds (if any) (no matter what it is).*

Observe that if we used *the pen John is looking for* as the antecedent for *it* in (B) we would have the referential reading all right: the pen's existence would be presupposed by the first sentence. However, when the first sentence is used non-referentially the pen can materialize (i.e., is available for reference) only when John eventually finds it. Refer also to Montague (1974f) and Heim (1982) for similar considerations. Finally, compare (A) and (B) with (C) below where the non-referential reading

for *it* as referring to the non-referential use of *a unicorn* is clearly perceptible.

(C)   *John is looking for* <u>a unicorn</u>.

      Bill wishes to see <u>it</u> *(the unicorn John finds)*.

      [because he is curious how a unicorn could look]

We wish to avoid any confusion when interpreting terms "referential use" or "non-referential use" as applied to some language expressions. By **referential use** of a (definite) description we mean the use where the speaker *intentionally* assumes or believes that there exists something which fits that description. In this sense the definite description in

(1)   *The present king of France is bald*

is interpreted **referentially** if we share such a belief. or just lack information to the contrary. The question whether the description was used to point to some particular individual or only attributively. Donnellan (1971). is not relevant at this stage and this is not the distinction we are making here. Later we will see that a part of what Donnellan (1971) called *attributive use* of a definite description is just a special case of non-referential (or rather semi-referential) use in attitude report contexts. The other classifications in the use of definite descriptions. for example. *inner attributive use. functional use.* Barwise & Perry (1983). or *generic use* belong to a different dimension than singular descriptions. and therefore must be treated differently; see Chapter 5.

Alternatively. when we use the term **non-referential use** we mean that the speaker *intentionally* does not refer to anything at all. That is. he knows or believes of nothing the description he uses is pointing to. or even if he does believe it refers to something. the existence of such a referent is not relevant to what he is saying. Thus

(2)   *John wants to marry a queen.*

      The queen must be wealthy.

with *the queen* used non-referentially. the speaker does not assume that anything being a queen actually exists. But unlike Donnellan's attributive use when we failed to refer to anything. thus causing our

statement be neither true nor false, see Donnellan (1971). here (2) may be true or false of John (as a part of his personal characteristic, for example).

4.4. Attitude Report Contexts

When a description is used consequently non-referentially the context setting situation changes so significantly that we need a separate rule to account for these cases. In Example 5 from section 4.3 when *a queen* was used referentially we would further describe her as *the queen John wants to marry*, while when interpreting the "story" non-referentially we could only speak of *the queen John (eventually) marries (if any)*. This difference has been properly accommodated by Rules 2 and 3.

We turn now to another class of verbs which can also create non-referential readings. These are **attitude report verbs**. Barwise & Perry (1983). like *imagine that*, *see that*, *believe that*, etc. We restrict our attention to the perfect contexts, in particular to those which do not involve any imperfect operators. The attitude report verbs used in such situations will be called **perfect attitude report verbs**. Although we are giving just one example with the attitude report verb *imagine that*, the discussion below applies to other perfect attitude report verbs as well.

Example 6

Let us analyze the following "story" in detail.

(6a) *John imagines that a unicorn lives in the park*

(6b) *The unicorn has a pink tail*.

Basically we can differentiate three reference situations between (6a) and (6b). As before we do not consider the trivial case where *a unicorn* of (6a) and *the unicorn* of (6b) do not co-refer. Assume first that we have the following translation as given:

* *unicorn* → *u*

  imagines that → *imt*

to have a pink tail $\rightarrow$ hpt

to live in the park $\rightarrow$ lp

CASE I. Suppose *a unicorn* in (6a) is used referentially, that is, its existence is presupposed. We obtain (we consider here just one possible referential reading)

- (6a) $\rightarrow$ $(\exists x\, (u\, x)\, \&\, (imt\, J\, (lp\, x)))$

  (6b) $\rightarrow$ $(\exists x\, (u\, x)\, \&\, (C\, x)\, \&\, (\forall y\, ((u\, y)\, \&\, (C\, y))\, \supset\, (x{=}y))\, \&\, (hpt\, x))$

  where the context C is $(\lambda x\, (u\, x)\, \&\, (imt\, J\, (lp\, x)))$.

according to Perfect Context Translation Rule (Rule 2).

CASE II. Suppose to the contrary that *a unicorn* in (6a) has been used non-referentially. That is

- (6a) $\rightarrow$ $(imt\, J\, (\exists x\, (u\, x)\, \&\, (lp\, x)))$

We have now two options for translating (6b). In one case *the unicorn* used in (6b) refers to some particular individual the speaker of (6b) knows but perhaps John does not. In this case

- (6b) $\rightarrow$ $(\exists x\, (u\, x)\, \&\, (C\, x)\, \&\, (\forall y\, ((u\, y)\, \&\, (C\, y))\, \supset\, (x{=}y))\, \&\, (hpt\, x))$

  where $C = (\lambda x\, (u\, x)\, \&\, (imt\, J\, (lp\, x)))$ as before.

This translation correctly emphasizes that from the point of view of the speaker of (6b) *a unicorn* in (6a) has been used referentially. Therefore **both** sentences get their referential translations, and case II reduces to case I. Observe that case II is that of misunderstanding the intention of the speaker, but it is how the hearer interprets the discourse at the moment. This situation may later be corrected to restore the original speaker's meaning if the assumed interpretation of discourse leads to an incoherent representation (see Chapter 6).

CASE III. Let us assume that the speaker of (6b) has the possibility of glancing into John's image of the unicorn and sees that it has a pink tail there. The speaker of (6b) is therefore extending our information of what John imagines to:

(6c) *John imagines that a unicorn that lives in the park has a pink tail*

Here, we should expect (6b) to translate as

- (6b) → *(imt J ($\exists x$ (u x) & (C x) & ($\forall y$ ((u y) & (C y)) $\supset$ (x=y)) & (hpt x)))*

    where $C = (\lambda x (imt J ((u x) & (lp x))))$

and can be read as

- *John imagines that the unicorn he imagines to live in the park has a pink tail*

The speaker of (6b) uses *the unicorn* semi-referentially, taking the image of unicorn as the context setting situation. It does not mean that the speaker of (6b) takes the image of the unicorn as the referent. The image is the only thing he knows about this unicorn, but he addresses the unicorn itself even if the latter does not exist. Moreover, the speaker of (6b) believes that the context image uniquely determines the unicorn, however non-existent, which enables him to use the definite descrip tion. He does not need any imperfect operator in his utterance. This is because the abstract situation created by an utterance of (6a) persists for some (short) period of time after the utterance took place. This situation involves an implicit export of an attitude report operator from (6a) into (6b). That is (6b) should be understood now as

- *John imagines that the unicorn has a pink tail*

but it is not necessary that the actants in (6a) and (6b) must co-refer. The attitude report verb *ima gine that*, although perfect (according to the definition given in section 4.3), has the ability to create non-referential readings, and once created, behave quite differently than an "ordinary" perfect verb (like those reported in section 4.2). One may wonder whether other attitude report verbs would behave in the same manner, especially "epistemic" attitudes such as *know that* or *see that*. It has been widely assumed that such attitudes are always referential, see, for example, Barwise and Perry (1983). It is important to avoid a confusion here. When α has the form of $(\exists x (P x) \& (Q x))$ then *know that*(α) does not say anything about the existence of x until we employ a general principle, known as Veridical-

ity Principle † and stating that *know that*$(\alpha) \supset \alpha$ Veridicality of some attitude report constructions does not change things a bit. If the principle applies in certain cases. it will help to simplify some expressions and reduce degree of ambiguity produced by the transformation. But before it is used. both translations. referential and non-referential. are legitimate and have to be considerent different.

This discovery calls for another translation rule. the **Attitude Report Context Translation Rule**

RULE 4 (Attitude Report Context Translation Rule)

An object *u* referenced by *the* has been recently mentioned in a *de dicto* environment with an attitude report verb **att** Let S_1, S_2 be defined as in Rule 2. Then the only non-referential reading of S_2 with respect to *u* can be obtained as

(i) $S_1(a\ P) \rightarrow (att\ (\ \exists u\ (P\ u)\ \&\ (F\ u)))$ ‡

(ii) $S_2(the\ P_1) \rightarrow (att_1\ (\ \exists u\ (C\ u)\ \&\ (\forall y\ ((P_1\ y)\ \&\ (C\ y))\ \supset\ (x=y))\ \&\ (P_1\ u)\ \&\ (F_1\ u)))$

The context C is derived from S_1 as $(\lambda u\ (att\ ((P\ u)\ \&\ (F\ u))))$ and att_1 is the implicit attitude report operator imported from S.

Thus far it appears that case III presents the only situation when a non-referential reading can be exported from an attitude report context. Montague (1974f) and Partee (1972) seem to agree with this observation. Otherwise we would always find ourselves in case II unless. perhaps. the sentence S_2 contained an imperfect construction. RULE 4 can be easily generalized over all cases where att_1 is an explicit attitude report construction in S_2 as in

*   *John believes that a unicorn resembles Mary.*

    He imagines that the animal has a single horn.

So stated. Rule 4 can also account for the situations where S_2 contains an imperfect construction imp_1 in place of att_1. The non-referential reading of the following paragraph where *the animal* is co-referred with *a unicorn* illustrates this point

- —*John believes that a unicorn resembles Mary*

  But the animal must have a horn.

An important outcome of Rule 4 in its original formulation is that it can explain a part of what Donnellan (1971) called the *attributive use* and Barwise and Perry (1983) called *value free use* of definite descriptions. Suppose someone says.

(1)   *The man drinking the martini is a fool.*

in the sense that the definite description *the man drinking the martini* is used attributively. Donnellan (1971). For a speaker to utter (1) is to explore an implicit context-setting situation in which there is *a man drinking the martini.* if the definite description used in (1) is to be singular. not generic. as we assume here. Therefore. to say (1) is to make the reference in the following context

(1a)   *I believe that there is a man drinking the martini.*

(1b)   *The man drinking the martini is a fool.*

The context-setting sentence (1a) is **implicit** here. A side-effect of this assumption is that one cannot use a pronoun in place of a definite description when saying (1b). In general any attitude report verb *att* can be used in (1a) and then imported to (1b) according to Rule 4. Three other examples illustrate this point further.

(2a)   *I imagine he has a wife.* (implicit)

(2b)   *His wife is the cook.* (explicit)

(3a)   *I think she has a husband.* (implicit)

(3b)   *Her husband is kind to her.* (explicit)

(4a) *I believe there is a book on my antique table.* (implicit)

(4b) *Take the book off my antique table!* (explicit)

We observe that. in part. the *attributive use* of singular definite descriptions is just a special case of the non-referential use in attitude report context. The latter does not account for all of Donnellan's attributive uses in particular for what Barwise and Perry named *inner attributive* and *functional* uses. We shall return to this problem in section 4.7 and also in Chapter 5 where we discuss the question of non-singular uses of definite descriptions.

4.4.1. Imagine

The structure and behavior of the imperfect verb *imagine* has a very interesting property not found in other imperfect verbs. As suggested in section 4 3. the perfected version of this verb may be verbalized as *have imagined (that)*. which suggests that the process of building some image has been completed. † Therefore in saying

(IM) *John imagines a unicorn*

we may actually communicate that

• *John tries to have imagined that a unicorn [exists]*

Except the straightforward perfect (referential) translation of

• *IM* → (|x (u x) & (im J x))

where *imagines* → *im*

one can obtain the imperfect (non-referential) reading of (IM) as

• *IM* → (try J (□x (imt J (u x))))

where *imt* is the translation of *have imagined that*.

---

† This version of "imagine" (with "that") can be compared to a belief. The other interpretation. of creating a mental image. is addressed in this section.

Indeed, *have imagined that* behaves exactly like the attitude report verb *imagine that*, ignoring the difference in tense. Now the Rule 4 allows us to derive yet another non-referential reading out of (IM). This is the imperfect attitude report non-referential reading

- $IM \rightarrow (try\ J\ (imt\ J\ (\exists x\ (u\ x))))$

Can we now combine *try* and *imt* in the above formula and replace them by *im*? In some sense we could do that, yielding the following translation:

- $IM \rightarrow (im\ J\ (\exists x\ (u\ x)))$

But now the air of imperfectness characteristic for the original utterance is lost, and in this reading *imagine* appears a perfect attitude report verb. Being both imperfect and an attitude report verb, *imagine* allows Rules 3 and 4 to be simultaneously applicable thus paving the way for four possible interpretations of an utterance containing it. As the reader has perhaps observed, the use of *imagine* as presented in this section, requires a real or alleged model of the image being created which is external to this image. Another use, named *creative use* involves creation of an image without a model whatsoever. We shall return to this issue later (Chapter 8).

We close, but of course do not exhaust, the discussion on inter-sentential references made by definite descriptions with the singular *the* determiner. As far as we have explored the problem, the translation THE3 suggested in section 4.2 proved itself correct and apropos when treating different reference situations. Our presentation would not be complete, however, if we did not mention of the references made by pronouns and briefly discuss the descriptive use of definite descriptions and their role in the information acquiring process. These topics are the subject of the next two sections.

4.5. Pronominal References

In a sense, a definite pronoun can be regarded as the most concise form of a definite anaphora. Although its referring capabilities are significantly narrowed as compared to a definite description, we

can formulate a set of translation rules for definite pronouns in discourse that closely parallel Rules 2.

3 and 4

Example 7

Consider the following "story".

(7a) *John wants$_1$ to catch a fish.*

(7b) *He$_0$ wants$_2$ to eat it$_1$.*

(Subscripts with the verb *want* are for identification purpose only.) The translation of pronouns has

been suggested by Montague (1974f) as

- $he_n \rightarrow (\lambda P\,(P\,x_n))$

where x_n is a free non-linguistic variable from the category e of names (see Chapter 2). We adopt this

representation here. Observe that according to this solution the variable x_n stands for a name (unk-

nown to us) of some individual we refer to. consciously or not. by the use of a pronoun. We do not

consider here cases where the pronoun *it* may be used in place of other sentence constituents than

names or individual definite descriptions Suppose further that we have

- $catch \rightarrow c$

  $eat \rightarrow e$

  $fish \rightarrow f$

Applying RULE 1 we easily obtain the translation

- $(7b) \rightarrow (w_2\,x_0\,(e\,x_0\,x_1))$

which correctly represents the meaning of (7b) if we have no idea to whom or to what the pronouns

he_0 and it_1 refer. We shall call such a representation **context-less** or **literal**. If. however, having (7a)

as a context setting sentence we decide that he_0 has been used to refer to *John*. we quickly modify

translation of (7b) along the lines of the following derivation.

- $(7b) \rightarrow$

  $[(\lambda P\ (P\ J)).\ (\lambda x\ (w_2\ x\ (e\ x\ x_1)))] \rightarrow$

  $((\lambda x\ (w_2\ x\ (e\ x\ x_1)))\ J) \rightarrow$

  $(w_2\ J\ (e\ J\ x_1))$

We feel this result is correct no matter what interpretation has been assigned to $(7a)$ based on the assumption that names are **rigid designators**. Kripke (1972). that is. they persistently refer to the same objects independently of the situation in which they are used. The same cannot be said. however. when we consider the reference situation between *a fish* from $(7a)$ and it_1 in $(7b)$.

Suppose the speaker of $(7a)$ used *a fish* referentially thus creating a *de re* translation of his utterance. By using *it* in $(7b)$ we may therefore refer to *the fish John wants to catch*. i.e.. in $(7b)$ we claim that there exists a particular fish John wants to catch and this fish (not any other) is to be eaten by him. The reference for it_1 will be therefore in the form:

- $(\lambda Q\ (\exists x\ (f\ x)\ \&\ (w_1\ J\ (c\ J\ x))\ \&\ (\forall y\ ((f\ y)\ \&\ (w_1\ J\ (c\ J\ y)))\ \supset\ (x=y))\ \&\ (Q\ x)))$

which when applied to our last translation of $(7b)$ (abstracted over x_1) will yield

- $(7b) \rightarrow$

  $(\exists x\ (f\ x)\ \&\ (w_1\ J\ (c\ J\ x))\ \&\ (\forall y\ ((f\ y)\ \&\ (w_1\ J\ (c\ J\ y)))\ \supset\ (x=y))\ \&\ (w_2\ J\ (e\ J\ x)))$

More generally. we can formulate the following rule.

RULE 5 (Perfect Pronominal Context Translation Rule)

When a context setting sentence S_1 has a referential reading in the form

(i) $S_1\ (a\ P) \rightarrow (\exists x\ (P\ x)\ \&\ (F\ x))$

and the object x is pronominally referenced by he_n in a sentence S_2 with literal meaning represented by

(ii) $S_2\ (he_n) \rightarrow (F_1\ x_n)$

then the translation of S_2 in context of S_1 is derived as

(iii) $S_2 (he_n) \rightarrow [C. (\lambda x (F_1 x))]$

where the context C is drawn as

(iv) $C = (\lambda Q (\exists x (P \; x) \; \& \; (F \; x) \; \& \; (\forall y ((P \; y) \; \& \; (F \; y)) \supset (x=y)) \; \& \; (Q \; x)))$. □

Observe that the rule applies also to the situations when he_n refers to a name, although in such a case the context C will be set by the name itself. References to proper names are discussed in section 4.6.

Suppose now that *a fish* in (7a) has been used non-referentially. What *it* of (7b) may refer to now is *the fish John catches (if any)*. That is, the catching will set the reference for *it*, see Partee (1972). The reference would now take the form

$(\lambda Q (\exists x (f \; x) \; \& \; (c \; J \; x) \; \& \; (\forall y ((f \; y) \; \& \; (c \; J \; y)) \supset (x=y)) \; \& \; (Q \; x)))$.

In this situation, the meaning of (7b) will be rather represented by

(7b) →

$(w_2 \; J \; (\exists x (f \; x) \; \& \; (c \; J \; x) \; \& \; (\forall y ((f \; y) \; \& \; (c \; J \; y)) \supset (x=y)) \; \& \; (e \; J \; x)))$.

The following rule summarizes this case formally.

RULE 6 (Imperfect Pronominal Context Translation Rule)

If a context setting sentence S_1 with imperfect operator **imp** is used to utter its non-referential reading, i.e.,

(i) $S_1 (a \; P) \rightarrow (imp (\exists x (P \; x) \; \& \; (F \; x)))$

and the object x has been further referenced by a pronoun he_n in a sentence S_2 with literal meaning of

(ii) $S_2 (he_n) \rightarrow (imp_1 (F_1 \; x_n))$

then the translation of S_2 in context of S_1 is obtained as the result of the following derivation

(iii) $S_2 (he_n) \rightarrow [(\lambda \Phi (imp_1 \; \Phi)). [C. (\lambda x (F_1 \; x))]]$

where the context C is drawn from S_1 as

(iv) $C = (\lambda Q (\exists x (P\ x)\ \&\ (F\ x)\ \&\ (\forall y ((P\ y)\ \&\ (F\ y))\ \supset\ (x=y))\ \&\ (Q\ x)))$. $\square$

It is possible to formulate further translation rules for other contextual situations in the spirit of Rules 5 and 6. By simple analogy the following rule may be expected for pronominal attitude report contexts.

RULE 7 (Attitude Report Pronominal Context Translation Rule)

If a context setting sentence S_1 with an attitude report verb **att** is used to utter its non-referential reading. i.e..

(i) $S_1 (a\ P)\ \rightarrow\ (\text{att}\ (\exists x\ (P\ x)\ \&\ (F\ x)))$

and the object x is further referenced by a pronoun he_n in a sentence S_2 with the literal meaning of

(ii) $S_2 (he_n)\ \rightarrow\ (\text{att}_1\ (F_1\ x_n))$

then the translation of S_2 in context of S_1 is obtained as the result of the following derivation

(iii) $[(\lambda \Phi\ (\text{att}_1\ \Phi))\ [C\ (\lambda x\ (F_1\ x))]]$

where the context C is drawn from S_1 as

(iv) $C = (\lambda Q (\exists x\ (\text{att}\ ((P\ x)\ \&\ (F\ x)))\ \&\ (\forall y\ (\text{att}\ ((P\ y)\ \&\ (F\ y)))\ \supset\ (x=y))\ \&\ (Q\ x)))$

and the attitude report operator of S_2 may be either imported or explicit. $\boxed{}$

Let us now turn to a different reference problem involving again the definite descriptions.

4.6. Referring to a Name

Unlike the examples from sections 4.3, 4.4 and 4.5, the definite description is not used to focus our attention on a particular object we are talking about but to extend the information on that object which is unique - because we have already known its name. One aspect of this problem has already

been discussed in section 4.3 (recall Example 2). and it influenced both the form of the formula THE3 and the rules of translation. especially Rules 2. 3 and 4. Consider the following example.

Example 8

Suppose one hears the following "story".

(8a) *Fatsy wants to catch a fish*.

(8b) *The cat wants to eat it$_0$*.

We shall concentrate here entirely on the possibility that the definite description *the cat* refers to the individual named *Fatsy* in context setting sentence (8a). Translation Rule 1 gives us the representation of (8a) as (concentrating on referential reading only):

- (8a) $\rightarrow$ $(\exists x\ (f\ x)\ \&\ (w\ F\ (c\ F\ x)))$

  where *Fatsy* $\rightarrow$ $(\lambda P\ (P\ F))$.

The literal *(de re)* reading of (8b) is given below.

- (8b) $\rightarrow$ $(\exists x\ (cat\ x)\ \&\ (C\ x)\ \&\ (\forall y\ ((cat\ y)\ \&\ (C\ y))\ \supset\ (x=y))\ \&\ (w\ x\ (e\ x\ x_0)))$

  where *it*$_0$ $\rightarrow$ $(\lambda P\ (P\ x_0))$ and *cat* $\rightarrow$ *cat*

To obtain context *C* for *the cat* as referring to *Fatsy* we abstract the translation of (8a) over its subject. thus getting

- $C = (\lambda s\ (\exists x\ (f\ x)\ \&\ (w\ s\ (c\ s\ x))))$.

We then supply *C* as an argument to the literal translation of (8b) above. obtaining an extended representation of (8b) as

- (8b) $\rightarrow$

  $(\exists x\ (cat\ x)\ \&\ (\exists u\ (f\ u)\ \&\ (w\ x\ (c\ x\ u)))\ \&$

  $(\forall y\ ((cat\ y)\ \&\ (\exists u\ (f\ u)\ \&\ (w\ y\ (c\ y\ u))))\ \supset\ (x=y))\ \&$

  $(w\ x\ (e\ x\ x_0)))$.

Thus far we have that *there exists a unique cat such that it wants to catch a fish and eath it$_0$*

We have merely applied Rule 2 to one of the possible referents for *the cat*. (We could easily co-index

the cat with *a fish)*. But we want to achieve more than just saying that the actant of (8*a*) wants to

eat *it$_0$*. We actually learn that this individual is nothing other than Fatsy. The name, as a rigid designa-

tor, gives us an unambiguous context, independent of circumstances. We therefore obtain:

- (8*b*) →

$$[(\lambda P \ (P \ F)). \ (\lambda x \ (cat \ x) \ \& \ (\exists u \ (f \ u) \ \& \ (w \ x \ (c \ x \ u))) \ \&$$

$$(\forall y \ ((cat \ y) \ \& \ (\exists u \ (f \ u) \ \& \ (w \ y \ (c \ y \ u)))) \supset (x{=}y)) \ \& \ (w \ x \ (e \ x \ x_0)))] \ \rightarrow$$

$$(cat \ F) \ \& \ (\exists u \ (f \ u) \ \& \ (w \ F \ (c \ F \ u))) \ \&$$

$$(\forall y \ ((cat \ y) \ \& \ (\exists u \ (f \ u) \ \& \ (w \ y \ (c \ y \ u)))) \supset (y{=}F)) \ \& \ (w \ F \ (e \ F \ x_0)).$$

The part under the universal quantifier reports that the only cat around that also wants to catch a fish

is Fatsy. This is non-trivial information about the situation described which cannot be derived from

literal translation of (8*b*). The next significant extension of our state of knowledge is the information

that Fatsy is *a cat*. Similar considerations may be given to pronominal references (as we have already

mentioned in association with Rule 5). The use of utterance attributes such as stress and intonation

will often decide whether a name is just contingent knowledge we acquire about an object we refer to

(we saw this situation in the last example). or whether a name is the ultimate referent which situation

will be discussed next. The approach we present allows for representing both cases. Let us consider a

further example.

Example 9

(9*a*) *Fatsy wants to catch a fish.*

(9*b*) *The cat belongs to John.*

The literal translation of (9b) can be easily obtained with formula THE3 as

- $9b \rightarrow (\exists x \ (cat \ x) \ \& \ (C \ x) \ \& \ (\forall y \ ((cat \ y) \ \& \ (C \ y)) \supset (x=y)) \ \& \ (bt \ x \ J))$

    where *belongs to* $\rightarrow$ *bt*.

Suppose that *the cat* in (9b) refers to *Fatsy* in (9a) as a particular individual known to the speaker of (9b) by name. (Let us call him *B* hence.) In this case we cannot take the situation described in (9a) as a context-setting situation. Rather, *B* refers to an individual whose uniqueness is beyond any doubt for himself. In other words. upon hearing the individual's name *B* can fetch an unique identifying context from his knowledge base. Observe that *B* cannot use *the cat* should he be aware of more than one *Fatsy* at the instant he utters his statement. *B* may. however. disambiguate his reference upon examining what the speaker of (9a) (call him *A*) has said. Assume though that the use of the definite article *the* in (9b) is the reaction of *B* when he hears the name Fatsy which. at least for *B*. unambiguously refers to some individual *B* knows of. The definite description used by *B* is entirely drawn from the context-setting situation he refers to. and which is now hidden from us. and perhaps the speaker *A* as well. This description conveys a piece of *B*'s state of knowledge or belief about the individual in question. and may vary from scanty remarks as

(9c) *It belongs to John*

to much more informative remarks like

(9d) *This awful animal belongs to John*

The fact that *B* exploits some external context-setting situation known to him becomes even more clear when *B* makes his reference mistakenly in the sense that *A* has had another individual in mind when uttering (9a). If *B*'s utterance does not clash with *A*'s knowledge base.then the latter acquires some false information of *Fatsy$_A$* which is actually of *Fatsy$_B$*. What *B* means by his utterance of (9b) is that

- *The individual F I refer to by **the cat** is a cat and its name is Fatsy. and this information is sufficient for me to pick up a unique individual. that is F. From my point of view the*

speaker A is talking of F too.

In other words the context-setting situation has the factual form of

- $C \rightarrow (\forall y \,((cat\ y) \,\&\, (Fatsy\ y)) \supset (y=F))$

The literal $(Fatsy\ y)$ which may be read as *being a Fatsy*. creates the core of the context-setting situation the speaker B is referring to. see Burge (1975). Barwise & Perry (1983). The other literal may or may not be relevant here depending on the way B reaches his unique referent (that is. whether B uses A's utterance for that or not). In general. therefore. the context C the speaker B uses in (9b) is expressed as

- $C = (\lambda x\ (Fatsy\ x))$

and the final translation of (9b) follows as

- $9b \rightarrow (cat\ F) \,\&\, (Fatsy\ F) \,\&\, (\forall y\,((cat\ y) \,\&\, (Fatsy\ y)) \supset (y=F)) \,\&\, (bt\ F\ J)$

Note that the use of the literal $(cat\ y)$ under the scope of the universal quantifier. although most probably significant for speaker B. may be redundant. or even invalid if it turns out that F is actually an alligator We have two possibilities in the latter case Either the audience B is addressing becomes misinformed. which can happen when they exploit different context-setting situations. or the audience can still pick up F correctly if the context-setting situation they are looking at contains F and the clause

- $(\forall y\ (Fatsy\ y) \supset (y=F))$

This time B cannot mislead his audience on the basis of the logical truth of the formula

- $(\phi \supset \psi) \supset (\phi \,\&\, P \supset \psi)$

where P stands for B's incorrect belief. $\square$

The argument above accounts for the nature of the value-loaded use of definite descriptions. see Barwise & Perry (1981) (or what Donnellan (1971) called referential). and can be applied in a similar fashion to other referential situations we have discussed.

We formalize this discussion with two new translation rules. The rules account for the two distinct situations mentioned which we have named **contingent** and **ultimate** references to a proper name.

RULE 8 (Names as Contingent Referents)

An individual N. named N. mentioned in a context-setting sentence S_1 of the form

(i) $\quad S_1(N) \rightarrow (F_1\ N)$

is further referenced in a sentence S_2 by a definite description *the P* referring to its action or description in S_1 rather than to its name. If the literal translation of S_2 is in the form

(ii) $\quad L \equiv (\exists x\ (P\ x)\ \&\ (C\ x)\ \&\ (\forall y\ ((P\ y)\ \&\ (C\ y)) \supset (x=y))\ \&\ (F_2\ x))$

then the translation of S_2 in context of S_1 is derived as

(iii) $\quad S_2(the\ P) \rightarrow [(\lambda p\ (p\ N)).\ (\lambda x\ [(\lambda C.L).\ (\lambda s\ (F_1\ s))])].\ \square$

RULE 9 (Names as Ultimate Referents)

An individual N. named N. mentioned by a sentence S_1 of the form

(i) $\quad S_1(N) \rightarrow (F_1\ N)$

is further referenced by *the P* in a sentence S_2 on the basis of its name only. If S_2 has the literal translation in the form

(ii) $\quad L \equiv (\exists x\ (P\ x)\ \&\ (C\ x)\ \&\ (\forall y\ ((P\ y)\ \&\ (C\ y)) \supset (x=y))\ \&\ (F_2\ x))$

then S_2 translates in the context of S_1 as

(iii) $\quad S_2(the\ P) \rightarrow [(\lambda p\ (p\ N)).\ (\lambda x\ [(\lambda C.L).\ (\lambda s\ (N\ s))])]$

where N is the predicative use of name N. ⊞

4.7. Conditional Contexts

In section 4.4 on attitude report contexts we found that a part of the *attributive use* of some nominal expressions could be explained in terms of non-referential attitude report readings. In these cases we assumed that the speaker was addressing an entity whose existence was relative to his attitude toward it (beliefs, imaginations, etc). But this examined just one side of the coin. In the following example we list just a few sentences where an attributive reading cannot be explained by attitudes.

<u>Example 10</u>

Apparently, the following sentences can be interpreted non-referentially without any reference to speaker attitudes.

(1)   *The man who is drinking the martini is a fool.*

(2)   *The cat that Mary buys is a burmese.*

(3)   *A man who kills somebody is a murderer.* □

It should be relatively clear that the sentences like (1), (2) and (3) above roughly fall under the following scheme

(4) $A/The \; \alpha \; such \; that \; P(him_k) \; F's$

or, in other words

(5)   **if** a/the α P's **then** he_k F's

Rewriting (5) more formally we obtain

(6)   **if** $(\exists x \, (\alpha \, x) \, \& \, (P \, x))$ **then** $(F \, him_k)$

Clearly (6) is just another way to express (4) if the latter is to be understood attributively. No such equivalence can be made when (4) is used referentially. The choice between *the* and *a* in (4) depends on speaker confidence as to the uniqueness of the entity established in the condition part of the sentence. Observe that in (6) the part between **if** and **then** constitutes our context-setting utterance S_1, and the part past **then** is our S_2. It is not necessary to assume a pronominal reference between S_1 and

S_2 as it has been suggested in (6). In fact the general form of a conditional context utterance may be taken as

(7)   if S_1 (a P) then S_2 (the P_1).

Although sentences like those of Example 10 translate into conditional context form, not all conditional context utterances can be expressed in terms similar to (1). (2) and (3) or the like, for the reasons that become clear shortly. The importance of conditional structures in natural language has been long acknowledged see only Webber (1979) and Heim (1982). In Webber's (1979) thesis on discourse anaphora, she recognizes both conditionals and attributives, but she did not see any correspondence between them. In this section we do not attempt to present a solution to the general problem of conditional constructions, which are widely discussed by Heim (1982). Instead, we point out that a certain class of singular attributive statements can be reduced to a certain class of conditional statements.

Example 11

Consider the following pairs of sentences. If the first sentence in a pair is used non-referentially, it has an equivalent conditional context reading expressed by the second sentence in the pair.

(a)   *The cat that Mary buys is a burmese*

      If Mary buys a cat, it is a burmese.

(b)   *The man who is drinking the martini is a fool*

      if a man is drinking the martini, he is a fool.

(c)   *The man who kills somebody is a murderer.*

      If a man kills somebody, he is a murderer.

Thus the first sentence in (a) gets its non-referential conditional context reading in the form:

(a_1) $(\exists x \, (cat \; x) \, \& \, (buys \; M \; x)) \supset (burmese \; x_n)$

Resolving pronominal reference of x_n to x we obtain

(a_2) $(\exists x \ (cat \ x) \ \& \ (buys \ M \ x)) \supset$

     $(\exists u \ (cat \ u) \ \& \ (buys \ M \ u) \ \& \ (\forall y \ ((cat \ y) \ \& \ (buys \ M \ y)) \supset (y=u)) \ \& \ (burmese \ u)) \ \Box$

This latter translation requires some explanation as it is important not to confuse things at this point. The translation is not to represent the meaning of the conditional statement featured in the second sentence of (a). when taken alone. Obviously. the conditional statement does not contain the uniqueness implication that is present in (a_2). see. for example. Heim (1982). This translation represents a singular. attributive reading of (a) in which the uniqueness implication is clear . What the first sentence in (a) says is that the definite noun phrase *the cat that Mary buys* has at most one antecendent. which may be either a particular cat. or a concept of such a cat (which interpretation we spare here) . If this sentence is used referentially then there is no problem with instantiating the referent which must be exactly one. When the sentence is used attributively. however. we cannot claim that the antecendent of the definite description actually exists. and therefore we cannot instantiate our reference. We can do that if the reference proves successful. that is, under the condition that the antecendent can be instantiated. Thus the paraphrase of the attributive statement in (a) we are aiming at is *If Mary buys a cat then there is only one cat she buys and this cat is a burmese*. Other conditional paraphrases in Example 11 should be understood in this way. even if this may not appear the most natural reading. as in (c). Note that if the uniqueness implication could not be passed from an attributive statement such as (a) to its conditional paraphrase. the translation featured in (a_2) would not be acceptable. This situation occurs if we interpret the first sentence in (a) as generic and the second sentence as addressing instances of a generic concept. here *the cat that Mary buys*. In such a case we can only go as far as (a_1). and then seek some more elaborate translation rule to resolve the internal anaphora. The interpretation of generics. and other non-singular concepts is discussed in Chapter 5.

The example above has been chosen so that both sides of the conditional context reading for (a),
(b) and (c). when taken alone, have a straightforward perfect referential reading. This reading does not
represent the general case. and other pronominal context translation rules may be useful for resolving
pronominal references between the sides - consider only *If John wants to marry a queen. she must
be wealthy* As we will see. this reading cannot be obtained from the imperfect reading of *The queen
John wants to marry must be wealthy*.

An interesting consequence of discovering the equivalence between purely attributive readings of
some sentences and their conditional readings is that we can now explain translations of utterances
previously considered generic. see Hirst (1983). Consider for example

- *A tiger is more dangerous than a cat.*

In a referential interpretation we talk of a particular tiger and some particular cat. In a conditional con-
text reading with respect to *a tiger*. for example we address some particular cat. but not a particular
tiger. In such a case we utter that *if there is a tiger. it is more dangerous that a (particular) cat*.
Still the conditional reading may be applied to both *a tiger* and *a cat* yielding something like: *if there
are a tiger and a cat the tiger is more dangerous than the cat*. I think the reader should not have
any problems with applying the principles suggested above to derive proper representations for each of
these readings However. the truly generic utterances cannot be verbalized this way. The following
sentence

- *The president is elected every four years*

when used at some "naming" level L_{+1} (to be explained in Chapter 5) cannot be equated with

- *If there is a president. he is elected every four years*

unless. of course. the latter is used in a generic sense too.

We summarize these observations as a formal translation rule: using the notion of *conditional
context pattern*. which (roughly speaking) has the form of $(\lambda S_1 (\lambda S_2 (S_1 \supset S_2))))$ as expressed in Λ.

RULE 10 (Conditional Context Translation Rule)

If a sentence S has a singular reading. referential over x translating to

(i) $S(\textbf{the/a } P) \rightarrow (\exists x\ (P\ x)\ \&\ (U\ x)\ \&\ (Q\ x))$.

and every referential singulat reading equivalent to it assumes the same form. where U is an optional uniqueness clause possibly present when *the* is used in S. i.e.

(ii) $U = (\lambda x\ (C\ x)\ \&\ (\forall y\ ((P\ y)\ \&\ (C\ y))\ \supset\ (x=y)))$

where C is an external context that *cannot* be instantiated. then the conditional context non-referential reading of S exists. and is obtained as the result of the following derivation

(iii) $S(\textbf{the/a } P) \rightarrow [[L.\ Q].\ P]$

where L is the conditional context pattern in the form

(iv) $L = (\lambda S_1\ (\lambda S_2\ ((\exists x\ (S_1\ x))\ \supset\ (S_2\ x_n))))$

and the pronominal reference of x_n is resolved to the argument of S_1 by one of the pronominal context translation rules. $\square$

Comments and Consequences

There seem to be numerous consequences of Rule 10. Some of them are quite striking

Consequent I. The literal U in the referential translation of S in Rule 10 is an important indication whether a referential reading of S is possible whatsoever. It contains the reference to a context-setting sentence S' possibly implicit to S. If no such context-setting sentence can be found. no referential reading exists. Consider

(S') *I met a man in the library.*

(S) *The guy I saw yesterday was a fool*

If *the guy ...* refers to *a man I met in the library* we have the referential reading. A weak reference †

---

† The notion of weak reference will be explained in Chapter 5. Informally. x weakly refers to a set Y if it refers to some y $\in Y$.

to S' can suffice too

(S') *I met some guys in the library*

(S) *A man I saw yesterday was a fool.*

If *a man* ... weakly refers to *some guys* ... the referential reading is assumed.

Consequent II. Condition (i) in the formulation of Rule 10 seems to be quite restrictive regarding the form a sentence can take in order to qualify for a conditional context non-referential reading. Observe that the literal P can be verbalized as

$\bullet$ $(\lambda x\ (P_1\ x)\ such\ that\ (Q_1\ x))$

as. for example. in *a man who breaks the law*. Yet the restrictive wh-clause is not strictly necessary to maintain P in the requested form. If one utters

$\bullet$ *A tiger is dangerous.*

a non-referential interpretation is still possible in conditional context with $P = an\ entity\ such\ that\ it$ *is a tiger.* so that P translates as

$\bullet$ $P \rightarrow (\lambda x\ (E\ x)\ \&\ (tiger\ x))$

where E is the *entity* predicate which as non-significant. may be omitted

The question naturally arises: What kind of sentences do not qualify for conditional context readings? Obviously. the utterance which presupposes the existence does not qualify. (But this very sentence does!) Compare

A   (i) *There is a unicorn with a pink tail which lives in the park.*

   (ii) *A unicorn which has a pink tail lives in the park.*

B   (i) *There is a man that killed somebody (and) who will be prosecuted.*

   (ii) *A man who killed somebody will be prosecuted.*

Both A(i) and B(i) have strictly referential readings, and no non-referential reading is ever possible.

Observe, however, that A(i) has an equivalent referential reading of the form

- $A(i) \rightarrow (\exists x\ (u-with-pink-tail-which-lp\ x))\dagger$

which violates the restriction that *every* equivalent referential reading must meet form (i) from Rule 10. In contrast, A(ii) can be translated as

- $A(ii) \rightarrow (\exists x\ (u\ x)\ \&\ (hpt\ x)\ \&\ (lp\ x))$

with $P = (\lambda x\ (u\ x)\ \&\ (hpt\ x))$ and $Q = (\lambda x\ (lp\ x))$.

Consequent III. The next observation to make in conjunction with Rule 10 is that sentence S does not necessarily have to be a top level clause. Again consider the sentence

C.   *John wants to marry a queen.*

The referential reading of (C) assumes the form of

- $C(ref) \rightarrow (\exists x\ (q\ x)\ \&\ (w\ J\ (m\ J\ x)))$

Rule 10 immediately provides us with the conditional context perfect non-referential reading as

- $C(ccper) \rightarrow (\exists x\ (q\ x)) \supset (w\ J\ (m\ J\ her_n))$

This translations reads *if there is a queen, John wants (to marry her)* By Rule 3 we can also derive the imperfect non-referential reading of (C).

- $C(imp) \rightarrow (w\ J\ (\exists x\ (q\ x)\ \&\ (m\ J\ x)))$

And applying Rule 10 to the embedded clause under the scope of the existential quantification we get the conditional context imperfect non-referential reading of C

- $C(ccimp) \rightarrow (w\ J\ ((\exists x\ (q\ x)) \supset (m\ J\ her_n)))$

---

† One might overcome restriction (i) in Rule 10 by inserting the non significant literal "entity" in place of missing P or Q component. That would lead to the conditional context reading roughly paraphrased as: "if there is something, it Q's". This reading is, however, unlikely enough to be taken seriously.

This latter reading may be verbalized as: *John wants (to marry the queen if there is a queen)*.

There is a subtle difference in meaning between C(ccper) and C(ccimp). The former states that unless a queen exists John's attitude toward her cannot be instantiated. i.e., her existence causes John's desire to become in effect. In the latter interpretation John's will or desire is independent of a queen existence. Rather, her existence will cause John to marry her, relative of course to his will. More precisely, in C(ccimp) John wants to reach the state in which a queen's very existence will entail John's marrying her.

Similar considerations can be applied to utterances with attitude report verbs. Thus for the sentence

D   *John believes that a unicorn lives in the park.*

four non-equivalent translations may be derived:

$D(ref) \rightarrow (\exists x\ (u\ x)\ \&\ (b\ J\ (lp\ x)))$

$D(ccper) \rightarrow (\exists x\ (u\ x)) \supset (b\ J\ (lp\ x_n))$

$D(att) \rightarrow (b\ J\ (\exists x\ (u\ x)\ \&\ (lp\ x)))$

$D(ccatt) \rightarrow (b\ J\ ((\exists x\ (u\ x)) \supset (lp\ x_n)))$

Consequent IV. Let us briefly examine how the conditional context utterances behave in discourse. Consider for example. the following paragraph.

E   (i) *If a man breaks the law. he will be prosecuted.*

     (ii) *He will be tried and sentenced.*

With E(i) and E(ii) translated respectively as

$E(i) \rightarrow (\exists x\ (m\ x)\ \&\ (bl\ x)) \supset (wp\ x_n)$

$E(ii)(literal) \rightarrow (wts\ x_k)$

where *bl. wp.* and *wts* stand for *breaks the law. will be prosecuted.* and *will be tried and sentenced* respectively.

we need to resolve the pronominal reference of x_k in (ii) to either x or x_n in (i). Assume here the reference is to x. Applying Rule 5 between the condition part of E(i) and E(ii) we can get the full translation of the latter as

- $E(ii) \rightarrow (\exists x \, (m \, x) \, \& \, (bl \, x) \, \& \, (U \, x) \, \& \, (wts \, x))$

where U is as stated in Rule 5. But now the uniqueness clause U cannot be instantiated since it refers to the condition (hence non-referential) part of E(i). Therefore Rule 10 must be applied to the above yielding the ultimate translation of E(ii) as

- $E(ii)(cc) \rightarrow (\exists x \, (m \, x) \, \& \, (bl \, x)) \supset (wts \, x_m)$

with the pronominal reference of x_m resolved internally. Observe that if the clause U in the referential reading of E(ii) could be instantiated we would be forced to admit that E(i) has had a referential reading too. This kind of backup adjustment throughout a discourse will create an important aspect of the transformation F_n.

Finally, note that, although in the above example, we assumed the external reference to address the condition part of a conditional utterance, it does not have to be the case in general, and the conclusion part can be addressed as well. In such a case the uniqueness clause generated by a pronominal translation rule could not be instantiated for much the same reason. The decision as to which part of the conditional utterance is being addressed by an external reference also belongs to the subsequent transformation of F_n.

4.8. Indirect References

By an **indirect reference** we mean the reference situation where the object referenced by a definite description does not explicitly occur in the context-setting sentence but its presence therein can be inferred. To become more precise we first introduce the concept of **individual knowledge base**.

For every individual α, and every moment of time t there is a knowledge base KB_α^t such that it contains all facts about the universe α knows, believes, disagrees about, etc. at time t. Instead of the time instant t we may talk of some space-time location l, but its temporal expanse should be small enough to keep the knowledge base in a fairly static condition. Suppose then that at some location l, an individual α, the addressee in some discourse situation, is to interpret some utterance $S_2(y)$ with an anaphoric expression y to the preceding discourse. Let $S_1(x)$ be the current context-setting sentence, and let x be an object referred to by the description x in S_1. Note that x does not have to correspond to any nominal construction in preceding discourse (noun phrase, relative clause, etc.), and may involve entire sentences, or sets of sentences. Consider, for example, the following paragraph where the description x is spread over a set of consecutive sentences resulting in the object x of *six washed, cored, and cut into pieces apples*.

(1) *John bought six apples. He washed and cored them. Then he cut them into pieces.*

Let further KB_α^l be α's knowledge base at the time $S_2(y)$ is uttered. If KB_α^l contains the fact xRy such that y is another object, and R is a binary relation, and y' is some not necessarily explicit description of y such that it can be added to $S_1(x)$ thus creating $S_1'(x, y')$, then y refers indirectly to x iff y refers directly to y'. Here is a typical example of indirect reference:

(2) a. *John bought a car* $[x = $ the car John bought$]$

   b. *The engine works well.*

$$[KB_{John}^b = \left\{ car(x) \rightarrow \exists y\ engine(y)\ \&\ has(x, y) \right\}$$

$$y = x's\ engine;\ R = has;$$

$$y' = \text{the engine of the car John bought}]$$

For additional examples the reader is referred to Brown & Yule (1984). One must be careful when selecting the relation R which, as we saw, is all important in indirect references. Certainly not any binary relation can be appropriate. In the fragment (2), for example, the relation *has not* would not

create an indirect reference. In fact, we need R to be a function such that $y = R(x)$.† This function is determined by the context S_1, the current discourse topic, and some other related phenomena to be discussed in Chapter 6. It should not be difficult to imagine that in discourse (2) the addressee may recall some other part of his KB, for example that of

- $\left| \exists y \ engine(y) \ \& \ no{-}car{-}has(y) \right|$.

so that the $R = has{-}not$ would be appropriate. This attitude is relative to the development of discourse prior the utterance (2a).

We do not fully investigate the problem of indirect references in this thesis, beyond the general statements given above. In Chapter 6, we compare this phenomenon to another reference situation called *lost links*. We observe that the problem of indirect references is a special case of the inter-sentential reference phenomenon that we have addressed in this Chapter. When we assume that the context-setting sentence S_1 has already been evaluated against an individual's knowledge base before he attempts an interpretation of S_2, then the general reference scheme of context-setting sentence/current sentence is maintained. The only difference between direct and indirect references lies in the way the context-setting sentence S_1 is being produced.

4.9. Forward References

When we introduced our reference scheme at the beginning of this Chapter, we did not constrain in any way the mutual spatio-temporal relationship between the context-setting sentence S_1 and the current utterance S_2. The only requirement was a relative physical proximity of the two so that they could be considered as fragments of some larger discourse. Avoiding unnecessary complications, we yielded to the natural limitation of written language which imposes explicit left-to-right ordering of words. In this sense, the context S_1 always preceded the sentence S_2, and this was especially explicit

---

† To obtain this functional relationship we have to choose R more carefully. Possible candidates are: "has as an engine", "is powered by", etc.

in

the examples we presented thus far. What we have considered as a notational convenience before, now needs a careful classification.

In practical applications of our theory we cannot ignore the fact that the language expressions can be produced and recognized one at a time, which imposes a linear temporal ordering on utterances. This observation leads to a quite natural division of all reference situations into **backward references** and **forward references**. Perhaps surprisingly, the actual temporal relationship of S_1 and S_2 does not have decisive impact in this matter. Clearly, if the context-setting sentence S_1 has been obtained from a discourse fragment that wholly preceded the current utterance S_2, references made from S_2 to S_1 have the backward character. Most of the examples presented in this thesis exhibit this characteristic, which is only partially a matter of convenience: backward references are far more common in discourse than forward connections. There are, however, situations where the context-setting sentence is not available at the time the utterance S_2 is produced, often deliberately withheld by the speaker and supplied at some later time. If the addressee reacts accordingly, postponing resolution of references in S_2 until S_1 is provided we have the situation where S_2 precedes S_1 (more precisely, the discourse fragment that S_1 derives from). In such a case we speak of forward reference of which the following example is characteristic:

(1) *He ran quickly down the stairway. Inspector John Flynn had his reasons to be in such a*
    hurry

Unfortunately, the backward/forward distinction is not always so clear cut. Often S_1 and S_2 will overlap arbitrarily so that we need some more definite method to decide whether a backward or a forward reference is taking place. Sometimes the distinction will not make much sense, however. Consider, for example, the situation in which the current utterance, which includes S_2 in it, is itself a part of the context S_1. For the sake of completeness, however, we attempt to formally define the notions of backward and forward references.

Let $S_1(x)$ be a context-setting sentence. and x be a description which **fully describes** some unique object x to be referenced by an expression y of the "current" utterance $S_2(y)$. † Taking into account the temporal ordering of language. we say that y **refers backward** to x iff the description x of the object x wholly temporarily precedes the utterance of y. In such a case we speak of a **backward** **(co-)reference** between y and x. Conversely. if x is not fully defined prior to the utterance of y. we shall speak of **forward (co-)reference** between y and x. Initially. this definition may seem overly com- plicated. Consider. however. the following paragraph in which the context-setting sentence S_1 is. derived from the utterances (2a) and (2c). with the intervening "current" utterance S_2 as (2b).

(2) a. *John bought some delicious coffee in the pub.*

   b. *He drank it back in his office.*

   c. *The cream he used was not very fresh and the coffee tasted rather bad.*

Clearly. the proper antecedent for *it* in (2b) is *the delicious coffee John bought with the not very fresh cream he added*. and not just *the delicious coffee he bought*. what would be the case if *it* referred backward into (2a). In fact the antecedent of *it* is not yet fully defined in (2a). and only after reading (2c) can one resolve the reference properly. In this sense. *it* refers forward. although neither (2a) nor (2c) contains an explicit antecedent. Observe also that the definite description *the coffee* in (2c) refers backward. and has the same antecedent as *it* in (2b)

A word of warning is necessary here. After reading (2a) and (2b). the reader. quite rightly. to the conclusion that the pronominal *it* refers anaphorically (hence backward) to *the delicious coffee John bought in the pub*. Later. upon reading (2c). he may change his earlier decision and relate both *it* from (2b) and *the coffee* from (2c) to the common antecedent previously described. This change in the reference of *it* involves some non-local processing that does not fit into our simple S_1-S_2 scheme. and therefore cannot be solved by the transformation F_{n-1}. Note that if the reader of (2) had employed the scheme independently to (2a) and (2b) first. and then to (2a,b) and (2c). he would have obtained

---

† By **full description** of an object x we mean here that x is ready for a *direct* reference from $S_2(y)$. Observe that the no tion is relative to S_2 and speaker's intention. Note also that the description x may come from discourse as well as from an individual's knowledge base.

distinct antecedents for *it* and *the coffee*, an obviously undesired result. For this reason, and some others to be explained later, the transformation F_{n-1} does not take us yet at the ultimate stratum in the Stratified Model. Another level is still necessary, and we shall discuss the problem in Chapter 6.

4.10. Discussion

There are many other types of references which we do not discuss in this thesis. Among these, the *one*—anaphora, clausal *it*, and elliptical constructions are perhaps the most prominent. See also Webber (1979), and Brown & Yule (1984) for more comprehensive lists. These problems require separate investigation which is not covered in this thesis. A considerable number of rules resembling these of 2 to 10 have to be added before the transformation F_{n-1} could be regarded as nearing a complete form. We limited our discussion to an in-depth investigation of a small number of inter-sentential reference cases since we believe that further rules, while undoubtly important, will not change the general image of the transformation F_{n-1} which we have presented thus far. As a matter of fact, the next chapter will introduce two more rules to F_{n-1}, Rule 11 and Rule 12. These rules actually add a new dimension to the phenomenon of inter-sentential references, and co-references in general, that has never been explicitly addressed before in a uniform-approach. We now summarize the characteristics of the transformation F_{n-1} in the Stratified Model.

The transformation F_{n-1} takes, from the level SLA_{n-2}, a representation of some actual discourse entered originally in the source language SL and produces the subsequent representation of the former at the level SLA_{n-1}. Let $\Delta = \{\delta_1, \cdots, \delta_s\}$ be such a discourse representation at SLA_{n-2}. Without sacrificing generality we may assume that Δ is finite and fully ordered, i.e., $\delta_i < \delta_{i+1}$ for $1 \leqslant i < s$, where $<$ is a precedence relation. Each $\delta_i \in \Delta$ is an image of some actual expression from SL in categorial grammar (and other early transformations), and in most cases δ_i will be classified into the syntactic category of sentences (t in CAT). Because CAT is known to introduce ambiguity, and also

because of ambiguity introduced by earlier transformations. a δ_i may not be a single representation (a parse tree) but would be rather a set of these. That is. $\delta_i = \{\delta_i^1. \quad . \delta_i^j\}$ such that $\delta_i^j \in F_{n-2}(\delta)$. $\delta \in SLA_{n-3}$. It would be more convenient. however. to regard the SLA_{n-2} representation of discourse not as a single set Δ. but as a set Δ^* of sets such that every $\Delta \in \Delta^*$ is an alternative discourse representation. and if $\Delta = \{\delta_1. \quad . \delta_s\}$ then every $\delta_i \in \Delta$ is a single p-marker in CAT. In other words. let Δ^* be the SLA_{n-2} representation of some actual discourse. We have

$$(1) \quad \Delta^* = \left\{\Delta \mid \Delta = \{\delta_1. \quad \cdots . \delta_s\}. \delta_i \in F_{n-2}(u_i). u_i \in SLA_{n-3}\right\}$$

where u_i is the discourse fragment corresponding to δ_i.

For each $\Delta \in \Delta^*$ the transformation F_{n-1} produces the representation $F_{n-1}(\Delta)$ at the level SLA_{n-1}. as a set of prototype discourse representations. Let D be such a discourse prototype produced by F_{n-1} at the level SLA_{n-1}. By cross-section D_c of D. $D_c = <D_c^p. D_c^f>$. we shall mean the discourse representation D with a distinguished point c and such that if $D \in F_{n-1}(\{\delta_1. \quad \delta_c. \delta_{c+1} \quad \delta_s\})$ then $D_c^p \in F_{n-1}(\{\delta_1. \quad . \delta_c\})$. and $D_c^f \in F_{n-1}(\{\delta_{c+1}. \quad \delta_s\})$ Let $S_2 = \delta_c$ be the current expression. of Δ. in the sense that the set $\{\delta_1. \quad . \delta_{c-1}\}$ has already been processed into the level SLA_{n-1} Then the sum $D_{c-1}^p \cup D_c^f = S_1$ is the context-setting sentence for S_2. It is not to say that D_c^f and D_{c-1}^p are completely independent. In fact. D_c^f will build not only on $\{\delta_{c+1}. \quad . \delta_s\}$ but also on the context provided by D_{c-1}^p. Here D_{c-1}^p is a possible discourse representation accumulated by processing the discourse part from δ_1 to δ_{c-1} into SLA_{n-1}. The other constituent of S_1. D_c^f. stands for a presupposed representation of the rest of the discourse following S_2. In practical application the structure D_c^f does not have to be complete. or may be even ignored. but its theoretical role in resolving forward refer-

ences cannot be neglected

The processing of the transformation F_{n-1} can be now formulated into the following recursive rule Let $<S_1, S_2>$ be the context-setting sentence/current utterance pair, with $S_2 = \delta_i$. Then

(2) $\quad F_{n-1}(<S_1, S_2>) = \left\{ D_i \in SLA_{n-1} \mid \exists d_i \in F_{n-1}(S_2) \text{ s.t } D_i = S_1 \bigcup d_i \right\}$

For every such D_i new S_1 and S_2 are derived as follows:

(3) $\quad S_2 \leftarrow \delta_{i+1}, \ i < s$

$\qquad S_1 \leftarrow D_i' \bigcup D_{i+1}'$

This scheme introduces a considerable degree of ambiguity. The unconstrained F_{n-1} has to consider and process every pair $<S_1, S_2>$ that can emerge from using formulas (2) and (3), resulting in a combinatorial explosion of possibilities (highly undesirable in a practical implementation). The process is finite, however. There are only finitely many different S_1's that can ever be produced, and the number of different S_2's is exactly the same as the cardinality of Δ which is finite by definition. The process starts with $S_2 = \delta_1$ the empty D_0, and a number of alternative D_1' which depends on S_2. The process stops when for every discourse prototype D if $D_i \in SLA_{n-1}$ then $i = s$.

The complexity of this scheme can be reduced if we allow another transformation, F_n, to parallel F_{n-1}. As one may expect unrestricted F_{n-1} can produce a great number of discourse prototypes of which only a handful will have actual importance. The rest can perhaps be excluded as inconsistent, senseless etc. but the local character of the F_{n-1} processing does not give any guarantee that the proper prototypes will be selected. That is why we relegated this job entirely to the next transformation F_n which is outlined in Chapter 6. We have concentrated here on the performance of F_{n-1} on any particular pair $<S_1, S_2>$. In the following chapter we continue the description of F_{n-1} along similar lines

Chapter 5

Names and Descriptions

5.1. Non-singular Terms

Our account thus far provides the means to deal with different kinds of singular descriptions appearing in natural language utterances for which we hope to build a uniform strategy of translation into the lambda-calculus based meaning representation language Λ. We began with an idealized although non-trivial. subset of English FMT and showed that using Rules 1 through to 10 we can transform this subset of English from the representation level SLA_{n-2} into the representation level SLA_{n-1} according to our Theory of Stratified Meaning Representation outlined in Chapter 3. We use the adjective *idealized* when referring to FMT to mean at least three things. First and what is prob ably the most obvious. we apparently ignored many important aspects of "naturality" of natural language such as tense. modality. fuzziness. etc. This omission was necessary however to uncover and properly capture the nature of some other phenomena in the language usage. in particular that of making references in discourse. We draw an analogy to other scientific inquiry e.g. when physicists assumed. for a similar reason. the existence of the ideal gas or liquid. or point masses in Newtonian mechanics. The second kind of idealization derives from the fact that we assumed the translation rules apply to the utterances somehow linguistically preprocessed. so that. again in the range of that selected subset. Rule 1 may always be safely used. But this should not be considered as a flaw in the theory. We delegate the job of delivering suitable forms of original utterances to some early transformations F_j $(1 \leq j < n-1)$ in our stratified model. so that at level $SLA_{(n-2)}$. Rules 1 through 10 can form a piece of the transformation F_{n-1} which. in turn. can carry FMT onto level SLA_{n-1}.

Furthermore, we formulated our rules to deal with *singular* terms; singular descriptions, singular pronouns and names denoting singular objects. Although we appreciate the importance of the discourse situations involving non-singular terms, we restrained ourselves from appending appropriate rules to the transformation F_{n-1} before we are in position to understand properly their role in the language. The following Theory of Names and Descriptions accounts for the use of names and descriptions in natural language.

5.2. The Theory of Names and Descriptions

Let us consider some example sentences which appear to escape any treatment in our theory of translation thus far developed.

Example 1

Many philosophers and logicians. Quine (1960). Kripke (1972). Donnellan (1971). Vendler (1971). Montague (1974). Thomason (1980) Carlson (1982). Barwise & Perry (1983). appreciate that the usage of the underlined nominal phrases in the following sentences has a "general" or "generic" character except for "regular" singular interpretations which are possible only in some cases

(1a) The king *wears a crown*

(1b) The president *is elected every four years*

(1c) Gold *is a yellow metal*

(1d) Temperature *is a measure of molecular motion.*

One can imagine hundreds of similar examples involving such non-singular objects as *water, heat, the Pope, the number* etc. Certainly, the use any of Rules 2 to 10 does not suffice to properly capture their meaning. The application of formula THE3 to translate definite articles also makes no sense here. What we may hope to accomplish is only to derive the singular interpretations (if they exist at all).

Apparently. these terms belong to some other dimension than their singular counterparts

Unfortunately: there is no commonly accepted account of these species in philosophical litera- ture. Some authors. see Vendler (1971). Barwise & Perry (1983). cautiously called them *generic*. or *general* (for example *the king*). or *functional* (such as *the number of students. the temperature*) uses of (definite) descriptions .Others. like Kripke (1972). were quite close to considering them names (or at least some of them: *heat. gold*). Yet others. see Quine (1960) and (1973). advocate the notion of *abstract terms* as being made of attributes. such as *[being] red* (further abstracted as *redness*). or *[being] the man drinking the martini* (which cannot be so easily nominalized). which can predi- cate about "concrete" objects In this light. the so-called *attributive use* of singular definite descrip- tions as identified by Donnellan (1971). may be considered as addressing to some abstract. higher- level and therefore non-singular (in our interpretation) concepts Carlson (1982) discusses the case of the so-called *natural kinds*. a specific type among generic terms He advocates the view in which gen- eric terms are taken as denoting entities in the same way as ordinary singular terms do. This is an interesting view. with which we essentially agree though the approach remains deeply set into the possible world paradigm. see also Carlson (1978)

Quine (1960) presents the fullest discussion of various categories of terms found among natural language expressions In his account. almost everything one can say is made up of different kinds of terms (appropriately connected to yield meaningful utterances) which he classifies as singular. general relative abstract attributive. etc In our theory we do not classify terms so finely. although it is possi- ble that a future extension to the theory may develop along the lines approximating Quine's approach in some aspects. Before we expose our theory. however. we discuss a few examples which provide additional insight into our account of names and descriptions.

There are numerous striking linguistic puzzles involving definite descriptions which caused a headache for many theorists of language. Partee (1972). Montague (1974). Barwise and Perry (1983) are among those offering examples and comments. The following example illustrates the phenomenon.

Example 2

Consider the following inferences

(2a) *The temperature is rising.*

   The temperature is ninety.

so. *Ninety is rising.*

(2b) *The president met the Soviet leader many times.*

   The president is Reagan.

so. *Reagan met the Soviet leader many times.*

(2c) *The tiger lives in the jungle.*

   My pet is a tiger.

so. *My pet lives in the jungle.*

(2d) *The president lives in the White House*

   The president is Reagan

so *Reagan lives in the White House.*

The conclusions in (2a) to (2d) are obviously wrong. The explanation given by numerous researchers chiefly amounted to the corroboration that the definite descriptions *the temperature, the president* and *the tiger* in the first sentences of (2a), (2b), (2c) and (2d) respectively should be interpreted functionally i.e. as intensions. Montague (1974f). or functions over situations. Barwise & Perry (1983). Observe that if the descriptions were to be interpreted singularly or as enumerating all instances of a non-singular object (i.e. statements containing them were understood as making claims about each instance). the reasoning would be sound. The reason that examples (2a) to (2c) appear odd while example (2d) seems fine lies in our reluctance to apply singular interpretations in these contexts. Even

if we are inclined to say in (2d). that the first sentence should be understood as *every president lives in the White House*. the conclusion is incorrect in general terms. We claim that unless some two descriptions (or names) are used **singularly** or **measurably singularly** at the **same level** no simple reference can be made between them. In fact. another type of reference that we call **remote reference** can still take place and we shall put this view forward in this chapter. For now we can say only that inferences like those of. (2a) to (2d) are all false.

What does it mean for a description (or name) to be used **singularly**. or **measurably singularly**? What does it mean for two descriptions to be used **at the same level**? The answers to these questions lead us to the theory of names and descriptions which we explicate below.

Initially let us observe that our language deals with singular objects only. no matter how complex their structure happens to be. Suppose somebody is being put into the position of the Observer. who perceives all these objects and has to use his language to describe them. Some objects are sharply distinguished from others so he chooses to give them names as *John. Mary. Fatsy. Sun*. ... The others have no clearly perceivable boundaries but he still may name them: *tea. water. grass. snow*. .... and then refer to some measurable quantities of them as *some tea. little snow*. etc. Yet others appear to be numerous though enumerable. displaying strong similarities to one another. It would be pointless for Observer to give them each a name. Instead. he decides to refer to them as *a cow. the man. this tree*.† etc. Still. he prefers to say *the sun* or *the lake* rather than to invent new names if he is not sure how many of them are there. even if he is aware of just one specimen. Later he may find out that some objects were given identical names. so having encountered them together he must refer to one as *the John. the Sun*. or *a Fatsy*. Having completed his job. Observer. who is also a part of this world. may name himself *Observer* or *the Observer*. and happily sit down under *a tree* on *the grass*.

Let us call the whole collection of objects he has just described as **the Observer level**. and use the symbol L_0 for it. Suppose then we ask Observer to tell us as much as he can about L_0. Soon he

---

† One could well imagine this process as running top-town (from concept to an instance) rather than bottom up as suggested here. For the sake of argument. we ignore this other possibility at the moment. but we return to the issue later in this chapter.

finds out that his naming has its limits. As he discovers new facts about his world it becomes more and more cumbersome for him to communicate in terms of *every man, some cats, several trees, each president,* etc. He discovers that some things he originally considered distinct appear to be instances of some single object. Also he must admit that the identity of some other objects has to be put into question. Being smart enough, Observer invents two new levels. L_{+1} and L_{-1}, which augment his world. At level L_{+1} he places the new objects he discovered to be generalizations (or abstractions, if you like) of some measurable amount of objects from L_0 which displayed a striking similarity or even identity. From the perspective of L_{+1} he is able to tell us that *The tiger lives in the jungle,* that *The president is elected every four years,* and that *The Morning Star and The Evening Star are actually two appearances of the planet Venus.* The objects at L_{+1} are singular there, but they appear "generic" or "functional" or whatever of that sort as seen from L_0. Observe that these objects may not have straightforward measurably singular descriptions at L_0 (like *every tiger, some president,* etc.), and often it will not be possible to refer to them in the terms of the language available at L_0. In either case one may expect that some undescribable aspect of an L_{+1} object can emerge at L_0 even if they all have been derived from L_0 (which does not have to be the case). Next, Observer invents a new generation of names at L_{+1}. *the president* and *the tiger* may be among those names. On the other hand, Observer might prefer to use definite descriptions here, for the similar reason he frequently decided so at L_0. In fact, we have no means to distinguish between names and definite descriptions in discourse. We can only stick to linguistic conventions.

It probably would not take a long time before a new augmentation for L_{+1} becomes necessary. Two new levels L_{+1+1} and L_{+1-1} can be added in a much the same fashion. The level L_{+1-1} does not necessarily have to be L_0, although it probably will. More or less the same happens at the level L_{-1} where Observer can now say that what he previously considered to be *the atom* actually denotes many different kinds of atoms (H. O. Ca. Fe. etc.), that *tea* is not so uniform and many different teas can be found, and that under the name *Joe Alex* was actually hidden a group of crime story writers. Subsequently the level L_{-1} will expand by L_{-1+1} and L_{-1-1} with the former often different than L_0.

Let us now formalize our intuition.

Definition 1

A use of a description will be called **singular** if it denotes or refers to a singular object. A use of a description will be called **measurably singular** if it denotes or refers to some measurable quantity of a singular object. Otherwise we shall talk of **non-singular** use. ▣

It is difficult to give examples of these uses without a wider context. so we postpone examples until later. Observe that the definition does not exclude non-referential interpretations of descriptions.

Definition 2

A **level** will be an arbitrary collection of singular objects. A **level language** will contain these and only singular and measurably singular uses of descriptions communicating of the level objects. ▣

Definition 3

For any level L_n all names appearing in the L_n language have singular interpretations □

Definition 4

For any level L_n there will be at least two distinct levels L_{n-1} and L_{n+1} such that L_{n+1} contains the non-singular objects as seen from L_n. and L_{n-1} contains the objects for which the objects at L_n are non-singular. □

Definition 5

The Observer level L_0 is an arbitrary chosen level serving as a reference point. ▣

Although the Definition 4 is intuitively clear. especially in the context of the earlier discussion. it is still not obvious how one can get from some level L_0 to either L_{+1} or L_{-1}. First we shall show how a descent to L_{-1} can be done. and then how generalization leads to L_{+1}.

Suppose that we have an object N called N at level L_0. Let T be an arbitrary set we shall refer to as a *coordinate*. We shall often think of coordinates as ordered sets of which the time coordinate will be the most frequently used. In general, however, a coordinate does not have to be totally ordered (pure time coordinate), or even partially ordered (mixed space-time coordinates). One can imagine coordinates as arbitrary sets of "points" or "locations" at which certain general (or abstract) objects (for example: *the president, the atom*) are assigned more specific "extensions" or instances (such as *the president Reagan* or $H, Fe, Ca, ...$). It should also be quite clear that almost any object we can think of can be considered an instance of at least one more general concept, and often there will be more such concepts available. Consider, for example, water in a glass as being an instance of some totality of water in the universe (space coordinate), and as an instance of a concept of water such as in *water boils at 100 C*.

Suppose further that, for the coordinate T, the Observer discovers that the identity of N along that dimension can no longer be accepted. That is, there are at least two $x, y \in T$ such that N-at-x $\neq$ N-at-y. Without losing generality we can assume that the coordinate T has been chosen so that the following non-equation holds:

- $\forall x, y \in T. \; x \neq y. \; (N \; x) \neq (N \; y)$

Let $(N \; x)$ denote an object N_x for some $x \in T$. The Observer cannot place N_x's at L_0 without violating definitions 2 and 3. Instead he moves them onto a new level $L_{-1}^{N}{}^{T}$ leaving the original object N at L_0. N may be no longer a "real" object but the concept remains in language. $L_{-1}^{N}{}^{T}$ can be attached to any existing level provided that the definitions 1 to 4 will never be violated. It can also give a beginning to a new level. Note that the distribution of N over the coordinate T may force other objects X from L_0 to be distributed over T as well, and their instances placed at $L_{-1}^{X}{}^{T}$. This process may remain mostly implicit until we make an utterance relating $(N \; x)$ to other objects at $L_{-1}^{N}{}^{T}$. In general, we shall say that the level $L_{-1}^{N}{}^{T}$ is *lower* than the level L_0, and write $L_{-1}^{N}{}^{T} < L_0$. Often we shall drop the superscripts N and T over the level symbol assuming some lower level L_{-1} whenever it does not lead to

ambiguity. Observe that with the above account the level structure of objects has a dynamic, ever changing character. Any new empirical fact to be added to our world knowledge bears a potential reverberation in the level structure involving creation of new levels and moving objects between levels. At probably non-frequent idle states the definitions 1 to 4 assure the structure balance.

Moving at level L_{-1} the Observer is aware of an enumerable collection of different objects N_x's. Extending the description used for N over N_x's the Observer refers to them as *the N. a N. some N(s). every N*. etc. It is possible, of course, that some other object N' found at L_0 is now disclosed to be an N_x for some $x \in T$. What that means is that we have wrongly placed N' at L_0. because it actually belonged to L_{-1} (for example, *the Morning Star. the Evening Star*. and *the planet Venus*). But this was right at the time placed it. i.e.. it mirrored the state of our knowledge of the world at the time. We may now give names to some N_x's and N can very well happen among them. This time however N will not denote the old object from L_0: this will be actually quite a different name referring to one selected N_x. and which may be replaced by a definite description of $(N \; x)$. A schematic illustration of the descent process is depicted in Figure 5.1.

On the other hand. suppose we have some objects N_1. N_2. considered distinct at L_0. Suppose then that we discover some striking resemblance† between them along some dimension (coordinate) T. so that we need a generalizing concept to talk about them. We climb to some *higher* level $L_{+1}^{N,T}$. i.e. $L_0 < L_{+1}^{N,T}$. and establish a new object. a **superobject**. N there. Now as seen from $L_{+1}^{N,T}$ all N_i's are just the occurrences of N at L_0 at different values of coordinate T. In other words. the following equation holds:

- $\forall i \; \exists x. \; x \in T. \; (N \; x) = N_i$

Observe also that all N_i's now belong to the level $L_{+1-1}^{N,T}$ which is a part of L_0. As before we shall drop superscripts N and T for simplicity.

---

† This need not be any "ordinary" resemblance. and the objects may be collected upon an arbitrary key.
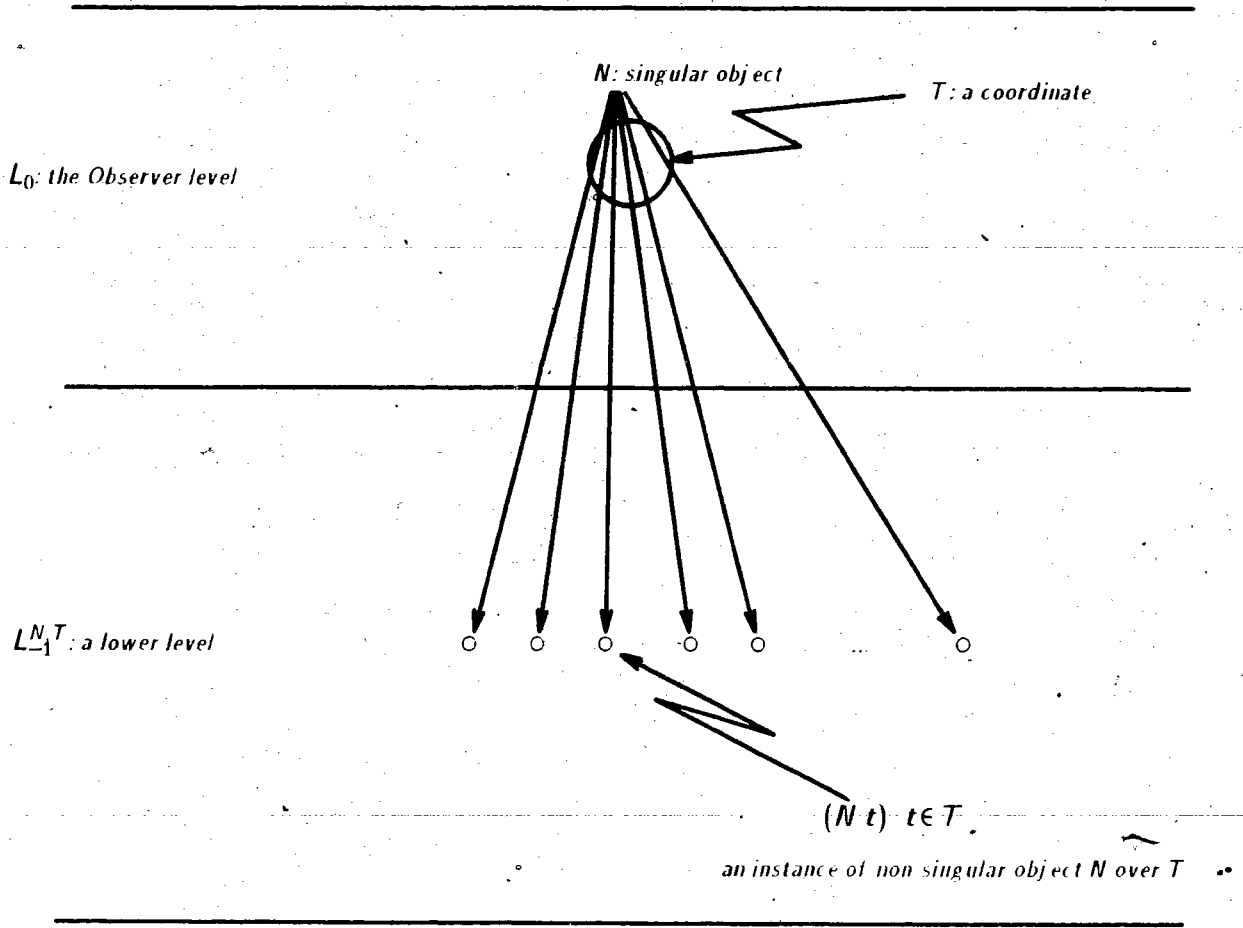
Figure 5.1. Descent from Observer level onto a lower level.

No matter how we name N at L_{+1} the following **Formula of Discovery** summarizes our action:

(FD) $\forall x \forall y. \; x.y \in T. \; (N \; x) = (N \; y)$

Remember that the formula FD is valid only when observed from L_{+1}. At L_0. N_i's remain distinct. traditionally - so they remain distinct in the language as well. If an object N satisfies formula FD with respect to some coordinate T we shall say that this object *maintains identity* with respect to (w.r.t) the coordinate T.

The generalization of other objects from L_0 onto L_{+1} may follow but, as in the case of decomposition discussed above, the process will remain largely implicit. Once the superobject N has been created it begins to live its own life. Some new objects from L_0, different than N_i's, may now become instances of N at some not yet utilized values of coordinate T. Also, we may use descriptions $(N\ x)$ without caring whether they actually refer to any objects at L_0. The latter property of general terms which is widely discussed by Quine (1960 and 1973) gets a formal explanation in our theory. It is important not to confuse a superobject with a set S of instances at L_0 which gave the beginning to this superobject with respect to a coordinate T. A superobject cannot be understood as a set of appropriate lower level instances, as we would obtain a measurably singular concept only. Instead, a superobject N can be identified with the function N from T into L_0 such that whenever $s \in S \subseteq L_0$ then there is $t \in T$ such that $(N\ t)=s$. The function N is then arbitrarily extended beyond the set S.

Computational tractability of such a function will depend then on what T is. In contrast to Montague's intensions of common nouns, we can concentrate on these aspects of generality of some concept which are of interest to us at the time. In particular, coordinates will often be just finite sets, or at most enumerable. There are, of course, infinite, continuous coordinates (such as pure time points coordinate). For example the relationship between the concept of temperature and the temperature at a certain point of time requires a continuous coordinate. Nonetheless, we may approximate the concept by some enumerable or even finite subset of time points at which the measuring is being done.

One further remark must be added here in order to avoid any misunderstanding. It is possible that some N_j from among N_i's was already recognized properly at L_0 as our goal object N from L_{+1}, although its other occurrences N_i for $i \neq j$ were not identified with it. In some sense, therefore, the previous concept of N was incomplete since it did not contain these other instances. The fact may be further stressed in our language when we choose to name N after N_j. This should not suggest that N and N_j are one and the same object. The former is somehow more mature although, when referenced by name, one can hardly tell which one of the two is being referred to, unless, of course, some additional clarifying context is present.

Let us now attempt to summarize the notion of a level in a somewhat more formal manner. We introduce the following relation of *relative singularity* between objects of the universe UA_{n-1} to complement the Definition 1 stated earlier in this chapter. The reader may observe that to guarantee transitivity of this relation we have to conceptually distinguish between instances of different objects (superobjects) even if these may prove identical in the original universe U.

<u>Definition 1A</u>

A object N is said to be singular relative to another object M (and write: $N R M$) iff both are instances of the same superobject w.r.t the same coordinate. i.e.. $\exists P.\ T\ \exists t_1.t_2 \in T\ \ P(t_1)=M\ \&\ P(t_2)=N.\ \square$

With the above assumption in force we can see that the relation R divides the universe of objects into mutually disjoint equivalence classes. These classes can be recognized as our naming sub-levels (such as $L^{N}_{-1}{}^{T}$.) It may be also observed that virtually any two objects can be classified as being relatively singular w.r.t some coordinates. while not w.r.t others. The fact that some two objects are not classified into the same class should not be interpreted that they remain in some instance-general concept relationship. The relation $<$. introduced earlier. imposes a partial ordering among levels which indicates where such relationships can be found

That. in rough terms. presents our Theory of Names and Descriptions This initial presentation should suffice for us to discuss some advantages we obtain from the theory when constructing the transformation F_{n-1} in the Stratified Model.

5.3. Remote References, Supercontexts and Subcontexts

Let us examine the impact the Theory of Names and Descriptions has on the transformation F_{n-1}. and how the account of inter-sentential dependencies presented in Chapter 4 can be extended over non-singular terms. We begin with a few examples.

Example 3

We have the following distinct objects at some level L_0: V called Venus. MS called Morning Star, and ES called Evening Star. Upon discovery that they all are just occurrences of the same planet we create a new object V′ named Venus at some level L_{+1}^{V}, ^{T} and such that for some $x.y.z \in T$, where T is a time coordinate. $(V′\ x) = V$. $(V′\ y) = MS$. $(V′\ z) = ES$. According to the FD formula we conclude from L_{+1} that $V=MS=ES$. while the same conclusion made at L_0 is false. □

Example 4

Let the level L_0 be as in the last example, except that the object V is discovered not to be uniform. In fact it contained occurrences of three different objects: planet Venus, and some two heavenly bodies assumed to be Venus in the mornings and the evenings. Now we cannot use our time coordinate T from the previous example to get the desired result of the object V′ at L_{+1}. Instead, we first descend to some L_{-1}^{V}, ^{S} over a coordinate S to differentiate the objects V_{s_1}, V_{s_2}, V_{s_3} for some s_1, s_2, $s_3 \in S$. Let the V_{s_1} be the part of the ultimate object V′. In the same way we create at L_{-1} all instances of MS and ES over the coordinate S. Let MS_{s_2} be called Morning Star at L_{-1} and ES_{s_3} be named Evening Star. Now we can construct the ultimate object V′ out of V_{s_1}, MS_{s_2} and ES_{s_3} over some other coordinate T (in fact it will be much like T from Example 3). We place V′ at L_{-1+1}^{V}, ^{T} which can give a beginning to some new level L_{-1+1}, or be just appended to L_0. Observe that V′ is singular at L_0.

This example is more "realistic" than the last one, but they both have equal linguistic significance. Note, however, that having L_{-1} as L_0 and L_{-1+1} as L_{+1} we would reconstruct Example 3. □

Example 5

At level L_0 the Observer is aware of the object T.P named The President. Let T be the time coordinate (different than in the last two examples). At L_0 we have, according to the FD formula that

- $\forall x \forall y. x.y \in T. (TP\ x) = (TP\ y)$

Later the Observer may discover that for some $t_1, t_2 \in T$, $(TP\ t_1)=N$ and $(TP\ t_2)=R$, and that at some level $L_1^{T\lceil T}$ where N and R belong, they are considered distinct and named Nixon and Reagan respectively. But at L_0, $R=N$ is true. The last observation can be made clearer if one imagines that TP is some abstract individual which (like Venus) when observed in early 70's is named Nixon, while when observed in 80's is called Reagan. □

The next two examples show how the theory contributes to our translation scheme.

Example 6

Consider the following paragraph.

(6a) *The president${}_1$ is elected every four years.*

(6b) *The president${}_2$ is Reagan.*

Except for our own common sense and intuition there is nothing in this "story" to prevent us from interpreting *the president${}_1$* and *the president${}_2$* at some levels L_1 and L_2 respectively so that one of the following takes place. Either $L_1=L_2$, or $L_1<L_2$, or $L_2<L_1$, or simply $L_1 \neq L_2$, where $<$ stands for the lower level relation introduced earlier. The latter case does not interest us since in such an interpretation both sentences were uttered at different occasions with no connection between them.

Suppose first that $L_1=L_2=L_0$. Then both sentences are translated with either Rule 2 or Rule 4 where the context setting sentences S_1^i and S_1' for (6a) and (6b) respectively are implicit and $S_1 \subseteq S_1$ only if the two definite descriptions are intended to co-refer. That interpretation, although possible, does not agree with our intuition. Observe that in this case the conclusion of

(6c) *Reagan is elected every four years.*

follows immediately

Assume then that $L_2=L_1^{T\lceil T}<L_1=L_0$ where TP is some object at L_1 and T is a coordinate. If *the president${}_1$* is used as a name we can expect the following translations

• $6a \rightarrow (eef y \ TP)$

$6b \rightarrow (\exists t \ (S \ t) \ \& \ ((TP \ t)=R))$

with *elected every four years* $\rightarrow$ *eefy. the president* $\rightarrow (\lambda P \ (P \ TP))$

where $t \in T$ and S is a selector over T provided by the discourse situation (for example *now. here.* etc.).

In a more general case we would take the phrase *the president*$_1$ as an ordinary definite description. Assuming some external context C we can translate (6a) with either Rule 2 or Rule 4. and the translation of (6b) will change accordingly. Assume Rule 2 is used in (6a). Then

• $6a \rightarrow (\exists r \ (p \ r) \ \& \ (C \ r) \ \& \ \forall y \ ((p \ y) \ \& \ (C \ y)) \supset (r=y)) \ \& \ (eef y \ r))$

$6b \rightarrow$

$\qquad (\exists x \ (p \ x) \ \& \ (C \ x) \ \& \ (\forall y \ ((p \ y) \ \& \ (C \ y)) \supset (x=y)) \ \&$

$\qquad (\exists t \ (S \ t) \ \& \ (p \ (x \ t)) \ \& \ ((x \ t)=R)))$

where *president* $\rightarrow \ p$.

The role of the literal $(p \ (x \ t))$ is significant in the situation where a description different from the one in the context-setting sentence is used to make the reference. for example. when stating some exception to a general rule. This point was already discussed in association with formulating translation for definite articles. and we won t repeat it here

A point that is worth discussion here is the role of the selector S in the translation of (6b). This selector may be interpreted as a local context at L_2 determined by a discourse situation. Barwise & Perry (1983). To visualize its influence on the form of the reference-making sentence compare the present example with that of (2c) repeated here for convenience.

(2c) *The tiger lives in the jungle.*

My pet is a tiger.

Unlike (2c). in the "story" of the president the local context allows us to use *the* in (6b) because there may be at most one instance of the general term at a given discourse situation (here the president of

the U.S.). This would not be the case had more than one value of the coordinate T satisfied the selector condition. The use of definite description in such a context suggests, therefore, that a single instance of a general term is being picked up by the selector. This means, in turn, that except for a main *remote reference* in cross-level referencing, a side local reference is being made at these levels where the definite description is used to denote some object. The remote reference itself does not need a definite description to be used for establishing a connection between the general term and its instance. The fact that we use a definite article when referring to an instance of a general concept (as in (6b)) implies only that the local context S contains (or is expected to contain) exactly one instance of the general object. Observe that the temporal aspect of this local context is extremely influential. †
When we decompose a general object with respect to a time coordinate, we are more likely to get a unique instance in a local context. However, when a coordinate does not contain a time element, as in (2c), we cannot, in general, exclude the possibility that instances other than the one we intend to refer to in subsequent statements may be present in a local context. Thus a definite description is not used until the local context gets properly narrowed. In the case of (2c) above, for example, we may continue the discourse at the lower level saying *The animal is quite friendly* and referring to the animal which is my pet tiger.

The above discussion lends a strong support to our Theory of Names and Descriptions, in particular to the existence of levels and coordinates. Having established a higher level object (or *superobject*) we can quite freely talk about its instances across various coordinates. The local references can be accounted for by the singular translation Rules 2 to 10.

At this point we are about to introduce a new translation rule. There are, however, two disturbing observations to be made about the translations above. These are the presence of some external context C, and the implicit assumption that the variable t changes over some not explicitly mentioned coordinate T. The former can be remedied instantly if we generalize the example in terms of sentences

---

. † This is, quite clearly, due to the temporal aspect of majority of sentences produced in natural language discourse (use of verb tenses and temporal adverbs).

S_1 and S_2 where S_1 sets the context. The latter will remain in the present form even if we can add some extra literals and a quantifier to the formula. Rather than unnecessarily complicating matters we suggest the clause $(\exists t \ (S \ t) \ \& \ ...)$ be read: *there exists a coordinate T such that there exists some value t of this coordinate such that* $t \in T \cap S$. *and* ...

Before Rule 11 can be formulated. however. we need to formally define the new concepts which have been introduced intuitively above.

Definition 6

An object **N** at a level L_n is said to be *remotely referenced* if the reference comes from some level L_m such that either $L_n < L_m$ or $L_m < L_n$. $\square$

Definition 7

A context in a remote reference will be called a *supercontext* if the reference comes from a lower level. The context will be called a *subcontext* if the reference comes from a higher level. $\square$

RULE 11 (Perfect Supercontext Translation Rule)

If the context-setting sentence S_1 with the translation

(i) $\quad S_1(P_1) \rightarrow (\ \exists x \ (P_1 \ x) \ \& \ (F_1 \ x))$

is interpreted at some level L_{+1}. and the object x is *remotely* referenced from the level L_0 by a description P_2 in a sentence S_2 with the literal translation of

(ii) $\quad L \equiv (\exists t \ (P_2 \ (u \ t)) \ \& \ (S \ t) \ \& \ (F_2 \ (u \ t))))$

then S_2 translates as

(iii) $\quad S_2(P_2) \rightarrow [C. \lambda u \ L]$

where the *supercontext* C is derived from S_1 as

(iv) $\quad C \equiv (\lambda Q \ (\exists x \ (P_1 \ x) \ \& \ (F_1 \ x) \ \& \ (\forall y \ ((P_1 \ y) \ \& \ (F_1 \ y)) \supset (x=y)) \ \& \ (Q \ x)))$

with S considered as a selector determined by the discourse situation. $\square$

Returning to our example, suppose now we have $L_0=L_1<L_2=L_{l+1}^{T,\mu,T}$. Our "story" may now look as follows

(6a) *A president$_1$ sits in the first row.*

(6b) *The president$_2$ is elected every four years.*

We have changed the first sentence for simplicity. The new (6b) has the following translation with a remote reference to *the president* in (6a):

● $6b \rightarrow$

   $(\exists x (p x) \& (\exists t (S t) \& (p (x t)) \& (sfr (x t))) \&$

   $(\forall y ((p y) \& (\exists t (S t) \& (sfr (y t)))) \supset (x=y)) \& (eef y x))$

   ·where *sits in the first row* $\rightarrow$ *sfr*.

The above formula says that the president an instance of which with respect to a point t of some coordinate T (space-time locations) sits in the first row is elected every four years. One may wonder whether the use of the definite description in (6b) is strictly necessary to maintain the remote reference. since we rejected such a necessity in supe-contexts. It is not easy to construct an example to support such a claim. We actually want to show that no such (non-artificial) example exists, and the use of definite descriptions in subcontext is forced by the level structure of the object world.

To support this contention let us observe that when we mention an object in the context-setting sentence intending it to be an instance of a higher level concept. we have already presupposed the uniqueness of such a superobject. and also have determined the coordinate along which the superobject is to be built. One can imagine, however. the object in context-setting sentence is an instance of different superobjects over perhaps different coordinates. so that it is not clear which superobject is being addressed. Nevertheless. when we address a superobject in a subsequent utterance it will not be clear whether any remote reference whatsoever is being made unless one of the following takes place:

either we restrict our attention to just these superobjects which our object is an instance of, or we invent some still higher-level concept generalizing over relevant superobjects. In the latter case the problem reduces to referring in supercontext. Consider for example the following

- *John wants to become a king rather than a president₁*

  That is because a president₃ is elected every several years, while the king rules for a lifetime

Here *a president₁* is an L_0 instance of an L_{+1} level object, say, *president₂* and *a president₃* is some still non-singular object (from the L_0 perspective) at some L_{+1-1} level. Obviously here, $L_0 \neq L_{+1-1}$. There is no direct correspondence between *a president₁* and *a president₃*. What relates these two "presidents" is a composition of two remote references the one in subcontext made by some L_{+1} object *president₂* and the other, in supercontext, made by *a president₃* to the *president₂* This example is a little artificial still, and one would prefer to say *the president* in the second utterance, or at least use some measurably singular description like *every president, some presidents,* or *most presidents* The objects and their respective levels just discussed are depicted graphically in Figure 5.2 •.

The former case is more complicated. Suppose we hear the following discourse

- *John serves as an editor for several scientific journals*

  ʃ *α editors are elected biannually.*

where α may be one of *some, some of these, all (of these), these* First, let us determine what kind of superobject John is intended to be an instance of. This superobject may be I: *scientific journal editor,* II: *editor of several scientific journals,* or simply, III: *editor.* In the case of I the level L_{+1} contains a collection of objects $E_1, E_2, \cdots E_n$ such that for some $j = 1, \cdots k, k \leq n, J = (E_{i}, t_j)$ for some coordinates $T_j, t_j \in T_j$. Let Σ be the set of all these E_{i}'s. In the case of II we have exactly one E_s such that $(E_s, t) = J$ for some $t \in T$. Finally, in the case of III, L_{+1} contains a single object E and John is just an instance of it over some coordinate T.
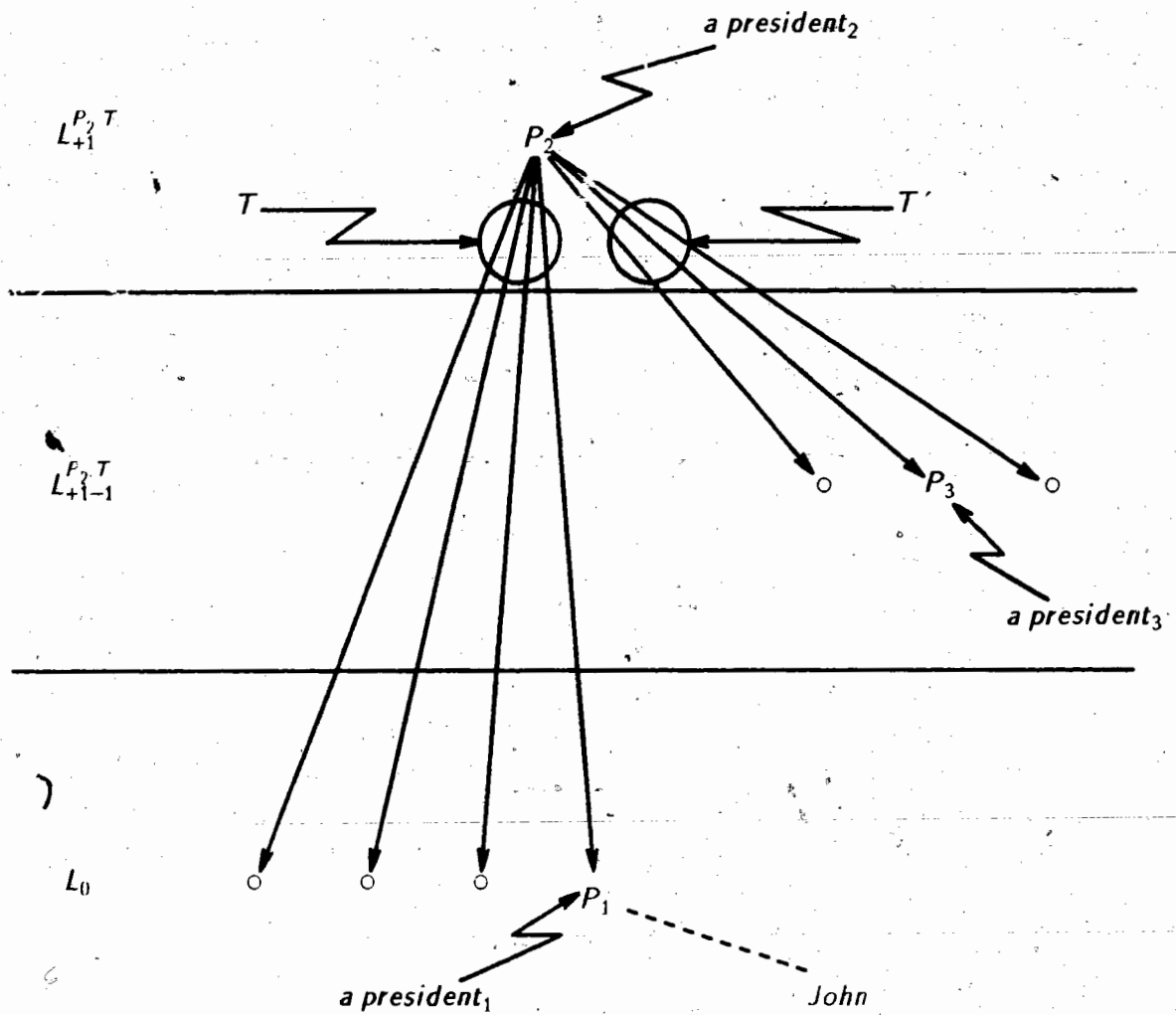
Figure 5.2. The 'president' example: another explanation

i.e., $(E\ t) = J$. The latter case obviously requires three levels to analyze the second sentence and is just another instance of the problem discussed above. In the two former cases when either *some* or *some of these* were used in place of α, we would refer to some subset $\Sigma' \subseteq \{E_1, \cdots E_m\}$ not necessarily being contained in the set Σ indicated by the first utterance, i.e. not necessarily $\Sigma' \subseteq \Sigma$. There-

fore. we cannot talk of any remote reference to take place. at least in the sense defined by our theory thus far.

A reading exists. especially with α representing *some of these* in the second sentence. where we are referring to some subset of these editors that John is an instance of. For this reading the description of *editors* or *these editors* refers exactly to the set Σ (case I). or (case II) to all instances of E_s over some other coordinate T_1 such that for any $t_1 \in T_1$. $(E_s t_1) = E_{s_{t_1}}$. and for some coordinate T_2 and $t_2 \in T_2$. $(E_{s_{t_1}} t_2) = J$. The superobject $E_{s_{t_1}}$ may be understood as an *editor of a scientific journal*. When we substitute *all (of these)* for α the situation is even more clear. A remote reference is actually being made. Because case II now reduces to case III. we shall concentrate further on the case I only.

Symbolically. we can represent the meaning of the first utterance as (see also Figure 5.3 for illustration):

- $\forall e \in \Sigma \; \exists t \in T \; such\; that\; (e\; t) = J$

Suppose *some* is taken for α. Then the second utterance translates as follows

- $\exists e' \in \Sigma \; such\; that\; (eb\; e')$

  where *elected biannually* $\rightarrow$ *eb*

Given a simple inference rule we can conclude therefore that

- $\exists t \in T \; such\; that\; (e'\; t) = J$

Here. however. the reference is made between a measurably singular object (the set of all these editors which John is an instance of) and some subset of it (these editors that are elected biannually). This is just a weaker version of a remote reference to a strictly singular object where the referent is a unique collection of objects instead of a single unique object. This phenomenon is not limited to remote references. and examples of "same level" references to measurably singular objects can easily be produced.†

---

† Consider for example: There are some apples on the table. Some of these apples are rotten.
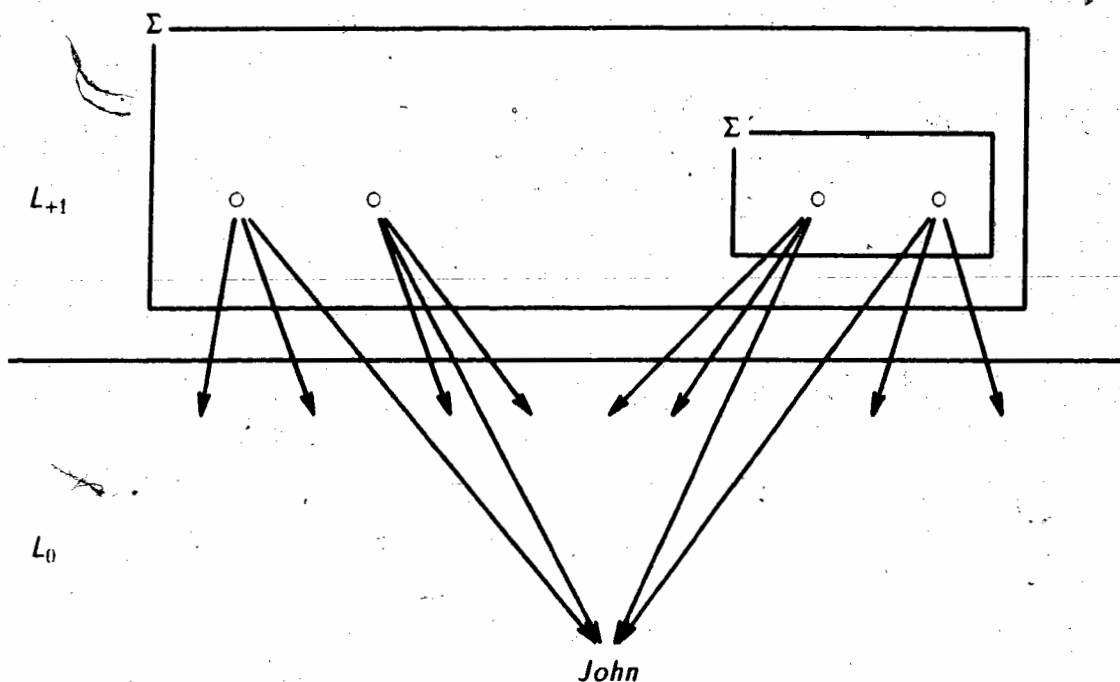
Figure 5.3. The 'editors' example (case I) weak remote reference.

There remains the last option for α which seems the most interesting. when α represents *these*. No matter what the scope of *these editors* happens to be. it would be now quite close for the second sentence to say

●   *The (Such an) editor is elected biannually.*

where *the editor* (or *such an editor*) generalizes over the scope of *these editors*. In fact we put forward the hypothesis that some measurably singular descriptions like *editors. these editors. dogs. presidents. tigers.* etc., somehow **asymptotically** approximate the superobjects *editor. dog. president. tiger.* etc. The more fine-grained a measurably singular description is, the closer it approaches the general singular term. The idea. found in numerous works. see for instance Quine (1973) and Dahl (1975). is often extended to the suggestion that the two are mutually exchangeable.

Measurably singular descriptions, no matter how fine-grained their referents happen to be, nevertheless express plurality, which cannot be said of general singular descriptions. We discuss this phenomenon further at the end of this chapter.

The following rule formally summarizes the case of remote references in subcontext.

RULE 12 (Perfect Subcontext Translation Rule)

If the context-setting sentence S_1 with the translation

(i) $S_1(P_1) \rightarrow (\exists x\ (P_1\ x)\ \&\ (F_1\ x))$

is interpreted at some level L_{-1}, and the object x is remotely referenced from the level L_0 by a definite description *the* P_2 in a sentence S_2 with the literal translation of

(ii) $L \equiv (\exists x\ (P_2\ x)\ \&\ (C\ x)\ \&\ (\forall y\ ((P_2\ y)\ \&\ (C\ y)) \supset (x=y))\ \&\ (F_2\ x)))$

then S_2 translates as

(iii) $S_2(P_2) \rightarrow [\lambda C.L.\ C]$

where the *subcontext* C is derived from S_1 as

(iv) $C \equiv (\lambda u\ (\ |t\ (S\ t)\ \&\ (P_1\ (u\ t))\ \&\ (F_1\ (u\ t)))$

with S considered as a selector determined by the discourse situation. $\square$

With Rules 11 and 12 we are getting a new fragment of the transformation F_{n-1} in the stratified model of meaning representation. These rules actually introduce a new dimension to F_{n-1}. While the Rules 2 through 10 operate within a single level only, the remote reference rules allow for crossing level boundaries. Remote references account for a class of inter-sentential dependencies in discourse that cannot be observed in the single-level model of the universe. We called the new rules *perfect;* however, other remote reference rules may be needed. It is possible that the other types of contexts (imperfect, attitude report, etc.) can be handled with either Rule 11 or Rule 12 where the context descriptions get adjusted to particular situations.

The use of pronouns in remote references may prove problematic. Although it appears perfectly right to use a personal pronoun in supercontext, it seems doubtful to make any sense in subcontext. Compare the "stories" below where the pronominal *he* is used in supercontext and then in subcontext.

- *The president is elected every four years.*

  He is Reagan now.

- *The president is sitting in the first row.*

  He is elected every four years.

The discussion of these problems is left for future research. Let us concentrate presently on another example which illustrates why, despite Rules 11 and 12 and examples (2a) to (2c) from Example 2, the stories like (2d) still seem to be logical and acceptable.

Example 7

Let us re-examine our reasoning of example (2d).

(7a) *The president lives in the White House.*

(7b) *The president is Reagan*

(7c) *Reagan lives in the White House.*

Does (7c) follow from (7a) and (7b)? The answer is *NO* when we use either Rule 11 or Rule 12 to translate the first two sentences. Observe that if *the president* in (7a) is a superobject with respect to that in (7b) it is not necessary that we make the claim about each instance. i.e., *the president usually lives in the White House.* Note also that this is not the problem of translation form, but the question of reference: if (7a) were interpreted as asserting about each instance of *the president* the translation would be reducible to an enumerably singular representation (with universal quantification), and the argument would be sound. The answer is *YES* if both (7a) and (7b) are interpreted at the same level (so is (7c)). and one of Rules 2 or 4 is used. The interpretation with Rule 2 is elementary. so let us concentrate on a possible use of Rule 4 here.

When Rule 4 is used to translate (7a) we assume the phrase *the president* there is used attributively (cf. Chapter 4). This rule gives us the following translation.

- $7a \rightarrow (att_1 (\exists x_1 (p\, x_1)\, \&\, (C_1\, x_1)\, \&\, (\forall y\, ((p\, y)\, \&\, (C_1\, y))\, \supset\, (x_1 = y))\, \&\, (lWH\, x_1)))$

  with *lives in the White House* $\rightarrow lWH$

Here, the attitude report operator att_1 is imported from some implicit context-setting sentence S_1
(7b) can translate with either Rule 2 or Rule 4 giving (as with Rule 2)

- $7b \rightarrow (\exists x_2 (p\, x_2)\, \&\, (C_2\, x_2)\, \&\, (\forall y\, (\quad)\, \&\, (x_2 = R))$

Now, if both (7a) and (7b) refer to the same context-setting situation S_1 then $C_1 = C_2$. and it follows
that $x_1 = x_2 = R$. Therefore, the conclusion in the form given below can be drawn

- $7c \rightarrow (att_3 (lWH\, R))$

Observe that the conclusion is relative to some attitude report operator. att_3 which was imposed by
att_1 of (7a) (and, perhaps, by some att_2 of (7b) if translated with Rule 4). This fact stresses the
attributive character of the above reference which in turn gives the illusion of a striking resemblance
to the examples which led us to Rules 11 and 12. It turns out however that the resemblance is
linguistic only, and that there is a deep semantic difference in our reasoning. □

5.4. Superobjects

Let us now examine the nature of *superobjects* i.e., the objects placed at level L_{+1}. We have
already put forward the hypothesis that the plural terms like *presidents, tigers, editors,* etc. asymptotically approximate superobjects *president, tiger, editor* respectively. As in the case of singular
superobjects, plural terms cannot be identified with sets of lower level instances over some coordinate
T. It turns out that the plural terms are actually prototypes of superobjects,† and they should therefore be placed at the same level as respective superobjects. We will see that the generalization leads

---

† Quine (1960) seems to strongly suggest this conclusion. Refer, for example to his double interpretation of mass terms
(pp. 120-121).

naturally to plural terms which may or may not induce equivalent singular superobjects. Conversely. a plural equivalent to a superobject may suggest the most natural coordinate to decompose the latter onto some lower level. When a superobject lacks a plural equivalent. however, we may admit that this superobject's origin has been traced down. A further decomposition is still possible but this process may often produce objects that will never assume an independent status and will remain recognized only as instances of this general concept scattered over that or another coordinate. This phenomenon is characteristic of the so-called *mass objects* and their corresponding *mass terms*. Quite naturally the question of where one level ends and another begins arises. The following example gives some insight into the problem of level boundaries.

Example 8

Consider the following sentences.

(8a) *Mary brings (some) water every day*

(8b) *John picks up the mail every morning*

(8c) *John buys bananas every weekend*

Let *water* in (8a) be the name of some superobject w at the level L_{+1}. Presumably Mary brings only a part of w but we can say that w is being brought by Mary every day. This is the same w every day. although each time possibly a different part of it is in transit. which leads to the obvious translation (at L_{+1})

(i) $8a \rightarrow (br-e-d\ M\ w)$

    where $br-e-d$ stands for *brings every day*.

On the other hand. suppose that Mary brings some water every day. Except for the above interpretation. we also have the measurably singular reading at L_0 where w is scattered over some coordinate T so that $\exists t \in T$ such that $(w\ t)$ is being brought by Mary. i.e..

$(\exists t\ (br\ M\ (w\ t)))$. This clause is. of course. relative to *every day* so at L_0 we could have

(ii) $8a \rightarrow (\forall x\, (d\, x) \supset (\exists t\, (br\; M\, (w\; t))))$ †

where $brings \rightarrow br$. $day \rightarrow d$

Both translations are essentially equivalent, and this equivalence is by no means accidental. For (8b) we can clearly produce the following interpretations.

(iii) $8b \rightarrow$

$(\exists x\, (ml\; x)\; \&\; (C\; x)\; \&\; (\forall y\, ((ml\; y)\; \&\; (C\; y)) \supset (x=y))\; \&\; (pick\text{-}e\text{-}m\; J\; x))$

where $picks\; up\; every\; morning \rightarrow pick\text{-}e\text{-}m$. $mail \rightarrow ml$. and the context C is used to distinguish John's mail.

Observe that x is an L_{+1} object in (iii) as well as in (iv) below which is suitable for the level L_0

(iv) $8b \rightarrow$

$(\exists x\, (ml\; x)\; \&\; (C\; x)\; \&\; (\forall y\, ((ml\; y)\; \&\; (C\; y)) \supset (x=y))\; \&$

$(\forall u\, (morning\; u) \supset (\exists t\, (pick\; J\, (x\; t)))))$

Similar translations may be expected for (8c). This time, however, the plural prototype will be used in place of the non-existent singular superobject. i.e.. $(buy\text{-}e\text{-}w\; J\; bnns)$ where $bnns$ stands for the plural term $bananas$. The reader may observe that the translation featured in (iv) can be reworded without a reference to the level L_{+1}. To do so we introduce a new predicate $John's\; mail$. and then rewrite (iv) as shown in (iv') below:

(iv') $(\forall u\, (morning\; u) \supset (\exists z\, (John's\text{-}mail\; z)\; \&\; (pick\; J\; z)))$

Are (iv) and (iv') equivalent? The answers is no: while (iv') is true only if for each morning we can instantiate the mail John obtains. (iv) remains true even if there are occasions that John gets no mail at all. Refer also to Carlson (1982) for similar considerations. A similar analysis can be applied to (8a) and (8c). Obviously the concept of a naming level extends the expressive power of our meaning representation language.

---

† To be precise we should represent Mary as $(M\; t)$. i.e.. as an instance of the L_{+1} object M at some $t \in T$. However. our naming convention discussed in section 5.2 allows for replacing the definite description by a new name at the level L_0. We utilize this option here.

There remains, however, one reading of (8b) which does not seem to require any reference to higher level objects. This reading could be paraphrased as *Every morning there exists the mail such that John picks it up*, or more formally

(v)   8b →

$(\forall x \, (morning \, x) \supset (\exists u \, (ml \, u) \, \& \, (C \, u) \, \&$

$(\forall y \, ((ml \, y) \, \& \, (C \, y)) \supset (u=y)) \, \& \, (pick \, J \, u))$

The variable u is apparently bound at L_0. Is this translation feasible? We can answer both yes and no. Yes, because the transformation gives us a singular interpretation of (8b) at L_0. Observe that because of the uniqueness clause in it, (v) says no more than that John keeps picking up the same thing every morning, since the context C does not depend on x and is the same each time. We can answer no because the latter interpretation most probably does not express our intention. Apparently therefore the translation of (v), although possible, is not equivalent to either (iii) or (iv). □

This discussion lends a strong support for our Theory of Names and Descriptions, and explains the intuition underlying its formulation (recall Example 6, which can be explained in the similar terms). The key observation of this section is the discovery of equivalence between plural terms and singular superterms (denoting superobjects). For now, we can assume that the plural terms cautiously named prototypes of superterms, actually denote superobjects as well. For some plural terms we can find equivalent singular versions like *tigers · tiger, presidents · president* etc. Others do not have this property, compare *bananas · books*. Alternatively the so-called *mass terms* will usually lack plural equivalents.[†] Still others may expose a surprising mixture of the properties, like *people* which is a plural term with morphologically singular form and which may also be used as mass superterm.

Let us, finally, in this chapter, analyze yet another example, quoted here from Hirst (1983), and show that a certain class of general statements of language which defied a satisfactory representation in other approaches finds an elegant solution in our theory. In the following sentence *computers*

[†] For example "water" cannot be identified with "waters" in general.

denotes a superobject.

> **IBM makes computers**

The L_{+1} translation is simple. $(mk\ IBM\ cmpts)$ where $makes \rightarrow mk$. But IBM probably makes (continually) some parts of the superobject $cmpts$. This leads us to an L_0 translation which necessarily involves decomposition of imperfect verb $makes$. The said translation is

- $(\tilde{mk}\ IBM\ (\exists t\ (\hat{mk}\ IBM\ (cmpts\ t))))$

  where $\tilde{mk}$ and $\hat{mk}$ stand for the imperfect operator (like try), and the perfect form (like $have\ made$) of the imperfect verb $make$.

5.5. Discussion

We briefly discuss how the Theory of Names and Descriptions contributes to the transformation F_{n-1} and the Stratified Model in general. Certainly, Rules 11 and 12 add a new dimension to the domain of text cohesion. They also introduce a great deal of ambiguity to the transformation F_{n-1}. Consider for example the context-setting sentence S_1. Every object x referenced in S_1, that we have previously considered uniform and atomic, can be now distributed over some almost arbitrarily selected coordinate or conversely may be thought of as an instance of some more general superobject over some other often arbitrary coordinate. With the notions of coordinate and superobject, the transformation F_{n-1} no longer has a discrete characteristic, which, at least in theory, may pose some problems for its computational realization. Fortunately, coordinates only rarely have to be specified explicitly. It will often be sufficient for our understanding of a remote reference, that the objects in question are defined at some different levels L_k and L_n such that either $L_k < L_n$ or $L_n < L_k$. That was the case with presidents, temperature, or water. Technically, nothing changed in our reference scheme. All of what we have said about the context-setting sentence, knowledge base, and current utterance remains in force. Moreover, we can regard remote references as a special, though distinguished case of indirect references defined in the last chapter. A superobject x with a coordinate T, available in the context-

setting sentence S_1 creates the function R which relates the superobject to one of its instances $(x\ t)$. $t \in T$ appearing in the current utterance S_2. For subcontexts. R is simply the membership function which for any instance $(x\ t)$ gives us the superobject x. Why, then, bother with remote refer ences at all?

There is one major consequence of the Theory of Names and Descriptions that may have an impact on the appearance of the stratum $<SLA_{n-1}. UA_{n-1}>$. and subsequently on the final stratum $<SLA_n. UA_n>$. including the mapping M In our discussion (to this point) we have deliberately avoided any indication of what one may expect at the universe side of the stratum $<SLA_{n-1}. UA_{n-1}>$. that is at UA_{n-1}. For lack of information to the contrary. and also in the light of the discussion in Chapter 2. the reader might quite understandably imagine UA_{n-1} as resembling the Montague's sys tem of possible worlds Nonetheless. the characteristics of the discourse meaning representation at the level SLA_{n-1} that emerged at the end of Chapter 4 suggested some smaller units than possible worlds to be recognized at UA_{n-1} In other words. if Δ was a discourse representation at the level SLA_{n-2} then $F_{n-1}(\Delta)$ appeared as a set of discourse prototypes: possible though not yet complete. discourse representations They were not complete in the sense that another transformation was necessary before we could attempt to define the semantics of such a representation Therefore. we would rather see UA_{n-1} as a set of possible "real situations" corresponding to the discourse prototypes at SLA_{n-1}

The Theory of Names and Descriptions takes us a step in this direction Observe that with the concept of superobject and coordinate we no longer have to identify general terms like *temperature* *president*. or *water* with intensions. i.e.. functions over possible worlds Superobjects are not some mysterious. extra-world entities. but they acquire a concrete status which makes them as comprehen sible as "ordinary" objects. To define an intensional term we no longer have to specify its extensions in all possible worlds what may require prior construction of these worlds Superobjects may be found in reality. and they are often quite material Can we then eliminate the possible worlds once and for all? It is much too early to provide a clear answer to this question Although the discussion in this chapter seems to suggest so. the Theory of Names and Descriptions in its present form addresses only a part

of the problem; more research is required to provide a sufficient evidence in this matter. But at least for the class of objects that have been customarily identified with the types e, $<s, e>$, $<s, <s, e>>$. etc in the possible world theory, we have managed to replace the notion of intension by the notions of superobject and coordinate defined locally, and not affecting explicitly other components of the universe. In a simplified account, our notion of coordinate can be loosely related to the possible world theory's concept of index (see section 2.1 for details). The coordinate is, however, far more selective than the index. In a particular discourse situation we can pick up that aspect of intensionality of some concept which is, at present, relevant to our understanding of the discourse. Our freedom in selecting that or another coordinate is all important. Note also that structures of coordinates may vary considerably. For example, a pure time coordinate may consist of time points as well as of time periods, and time periods are of much greater significance in practice (examine the *president*) But coordinates connote more than indices. They can be used to impose the non-singular status of an object which is otherwise purely extensional (the examples of *mail*, *water*, and *bananas* from the last section) In this sense, the concept of *non-singularity* has a local, and often subjective character Both objects and superobjects can coexist alongside in the universe, and as such may be elements of "real situations" at any UA_i, $1 \leq i \leq n$

We do not investigate the problems of deriving appropriate models UA_i of the universe in this thesis The discussion in this chapter should sufficiently depict the stratum $<SLA_{n-1}, UA_{n-1}>$. We now discuss some aspects of the next transformation F_n, possibly the last transformation on the Left Side of the Stratified Model This transformation is expected to deliver the ultimate discourse representation at the level SLA_n.

Chapter 6

Toward a Coherent Discourse Representation

6.1. An outline of the transformation F_n

In the last two chapters we considered the general problem of representing the meaning content of a natural language expression as uttered in the context of some abstract discourse situation. We made it explicit throughout those chapters that context is an important part of the interpretation of an utterance, and showed how context contributes to a sentence s meaning representation. We discussed solutions to a number of the more characteristic cases of inter sentential references created by the use of definite noun phrases and pronominal constructions. We also presented a dozen specific formal rules instructing how to compute the meaning representation for utterances set in well defined context situations, and with presumed connections to the surrounding discourse. The scheme can be roughly characterized as the following. Having an utterance $\phi_{i,\tau}$, a context c and a set τ of possible connections between the utterance and the context, find the representation $\phi'_{i,\tau}$ of $\phi_{i,\tau}$ assuming a particular interpretation c' of the context. The rules we formulated in Chapters 4 and 5 as well as those which existence was only speculated upon, conformed with this scheme creating the core of the transformation F_n in the Stratified Model.

There are two major sources of ambiguity arising from such a viewing of the transformation F_{n-1}. The first problem lies in establishing a unique interpretation of utterance context. To illustrate this point let us recall the concept of the context-setting sentence S_1 introduced in Chapter 4. Even with the set τ of cohesive ties determined, we often had several possibilities for reading the current utterance S_2, depending on the reading of S_1. This phenomenon has been explicitly captured by

providing a selection of rules for each cohesive relation considered between S_1 and S_2. We were virtually free to assume a translation of S_2 with either Rules 2. 3. or 4. or 5. 6. or 7. or even between Rules 8 and 9. Moreover. the addition of remote reference rules 11 and 12 to F_{n-1} significantly widened the scope of ambiguity in each case.

From a different perspective. having agreed upon some particular interpretation of the context, we can often observe a number of possibilities to connect S_2 and S_1 into a coherent fragment of discourse. Recall one of our earliest examples discussed in Chapter 4.

(1) S_1: *John interviewed a man.*

    S_2: *The bastard killed him.*

We have two (if no more) options to read S_2. Either we co-refer *the bastard* with *John* (using one of the rules 8 or 9). and *him* with *the man John interviewed* (using one of the pronominal rules 5 or 6). or we choose just conversely. In general. the larger the set τ. the greater (theoretically) the number of possibilities to relate the current utterance to the context.

In effect. when applying the transformation F_{n-1} to an utterance $\phi_{c, \tau}$ we cannot expect a single representation for ϕ but a collection of alternative representations relative to a particular interpretation of c and τ. That is. for some $\phi \in SLA_{n-2}$:

(2) $F_{n-1}(\phi_{c, \tau}) =$

$$\left\{ \phi_{c', \tau'} \in SLA_{n-1} \,|\, c' \text{ is an interpretation of } c. \text{ and } \tau' \text{ is an instantiation of } \tau \right\} \subseteq SLA_{n-1}$$

Suppose then we have a larger collection of utterances $\Delta \subseteq SLA_{n-2}$ constituting a hypothetic discourse. The translation $F_{n-1}(\Delta)$ of Δ into the stratum SLA_{n-1} will be a set D^* of **discourse prototypes** D such that:

(3) $$D^* = \left\{ D \,|\, \forall \phi \in \Delta \subseteq SLA_{n-2} \, \exists \phi_{c, \tau} \in SLA_{n-1} \; \phi_{c, \tau} \in D \right\}$$

A discourse prototype may not, in general, constitute a discourse yet. In spite of the starting hypothesis that Δ is connected and coherent, and therefore creates a discourse, the transformation F_{n-1} may still fail to account for some cohesive relations within Δ, thus producing only partially connected and partially ordered sets D. The question of ordering utterances within some D is not a trivial one, especially when we allow both forward and backward connections to context-setting sentences not immediately preceding or following the current utterance. For the discussion in this chapter, we assume that the ordering within D is uniquely superimposed by the physical ordering of the set Δ. This still leaves us with the problem of finding the "lost" cohesive links between subsets of a D which cannot be predicted by means of the co-reference relations computed locally by the transformation F_{n-1} (if such connections exist at all).

We now formulate the first approximation of what the transformation F_n is expected to solve in order to deliver the ultimate representation of a discourse fragment at the level SLA_n. If the collection Δ of utterances is indeed connected, and has a singular, coherent meaning, we would like the level SLA_n to be able to represent this meaning, perhaps in a form resembling that of an abstract situation, (Barwise & Perry 1983). If Δ has more than one possible meaning or reading that cannot be pruned by widening or deepening the discourse context (for example a literal and metaphoric reading) we would want all such readings to find their place at the level SLA_n. To guarantee this condition, at least partially, the transformation F_n must meet two major requirements. (A) From the set D^* of discourse prototypes provided at the level SLA_{n-1} the transformation F_n should select a subset $\hat{D}$ of possible candidates for the final discourse representation; and (B) provide the missing cohesive links within every D selected. In addition, we may expect F_n to discover and remove (if possible) any remaining observable ambiguity in every $D \in \hat{D}$. In this chapter we concentrate on problems (A) and (B) only. Summarizing, let D^* be a set of discourse prototypes derived at the stage SLA_{n-1} from an actual discourse Δ. Then

(4) $F_n(D^*) = \hat{D} =$

$$\left\{ \overline{D} \in SLA_n \mid \exists D \in D^* \; F_n(D) = \overline{D}, \text{ and } \overline{D} \text{ is } fully \text{ connected and } coherent \right\}$$

Observe that although F_n may still introduce some ambiguity so that $F_n(D)$ will be a set rather than a single discourse representation, we ignore this consideration. Note also that despite the sequential character of the transformations F_{n-2}, F_{n-1} and F_n suggested by the schematic illustration of the Stratified Model in Chapter 3, the latter transformation should, in practice, parallel the former two transformations.†

In the following sections we discuss the problems one must consider in order to construct a transformation F_n as outlined above. We do not propose any definite solutions, nor attempt to formulate pieces of the transformation. Instead, we indicate a number of the issues involved in constructing F_n, and estimate the range and difficulty of the problems we encounter.

6.2. The discourse phenomenon

A discourse is a communicative act involving a speaker or writer (the addresser), a message, spoken or written, expressed by means of some language, and an audience, perhaps a single hearer (the addressee) that is a recipient of this message. These characteristics of discourse are widely accepted in linguistic literature refer, for example to Brown & Yule (1984), Grosz (1977), Webber (1979), Grosz & Sidner (1985). Recall also the Situation Theory notion of discourse situation discussed in Chapter 2 (Barwise & Perry 1983). It should be clear that the three elements of a discourse situation need not occur in any kind of "physical" proximity, and this is especially true of written language. If such a proximity takes place, allowing for role switching between the speaker and his audience, we may talk of conversational discourse.

The first problem that any discourse analyst has to face is how to represent a discourse. Certainly, a full recording of the message communicated (a text) would be a part of such a representation. But the recording will not suffice in general. If we consider a particular discourse situation we must take into account what the speaker intended to put forward in his message, as well as how the hearer

---

† This point is of special significance when we consider a computer implementation of the Stratified Model, especially in respect to time and memory efficiency.

understood it. This. in turn. involves a set of expectations and presuppositions made by the speaker

about his audience. and by the audience about the speaker. The speaker may choose that or another

version of his intended message depending on what he believes the audience is likely to understand.

including a presumed audience's attitude toward himself. More or less the same happens for the

hearer. The audience tries to decode the speaker's message assuming some particular characteristics

of the speaker what again includes the audience's version of how he perceives him or them. This

approach leads naturally to a significant level of nesting of mutual presuppositions made by the parties

of discourse about one another. and for now we have no definite method for representing such situa-

tions. For that reason most of the contemporary research on discourse problems concentrated on

developing a simplified version of a **discourse model**. that is. a discourse representation as perceived

by one of the parties in discourse: either the speaker or the hearer. The concept of a discourse model

which originates from Webber (1979). can be found in related literature (although particular definitions

and ranges of the notion may differ) under the name of *focus. world model*. or *discourse represen-*

tation; refer. for example. to Grosz (1977). Cohen (1978). or Allen and Perrault (1980). There are two

important aspects of a discourse model which have to be clearly separated. One is the structural

organization of discourse into various kinds of *units* (Pollanyi & Scha 1984). or *segments* (Grosz &

Sidner 1985). reflecting those discourse fragments that exhibit some sort of internal consistency such

as syntactic form. common theme. level of detail. or common topic. These units may create fairly deep

stack-like structures representing various levels of *focusing* that dynamically changes as the discourse

proceeds: see. for example. Pollanyi and Scha (1984). Cohen (1984). or Grosz and Sidner (1985). Simi

larly to the structural programming. this organization of a discourse model can be used for dynamic

derivation of active context (see section 6.4) which. in turn. will play a significant role in selecting

proper antecedents for anaphoric. and other. references in text. The other aspect of a discourse model

is the underlying meaning content of the speaker's communication that actually decides on discourse

coherence. In the rest of this chapter we shall mostly disregard the structural organization of a

discourse model. and concentrate only on the meaning content it carries.

According to this view, the speaker maintains an encoded representation of some "real" situation he wishes, for some reason, to communicate to the hearer. What the speaker intends to achieve by this act is all important. If his intention is honest, the speaker wants the hearer to build up the same or similar model which the hearer could then, perhaps indirectly, decode into some other "real" situation. Ideally, it would be the same situation the speaker is talking about. But it may not always be possible, or even intended. We think that what the speaker wants to communicate is not a model of some part of the reality, or indirectly that part of the reality, although it may appear that way in many cases. What is really significant here is a sense of understanding of the communicated message which can be identified with some function that maps the discourse model onto some part of the reality. Suppose then that the speaker is talking about some "real" situation s_1 maintaining some representation s_1' of it (at some SLA_n level). What the speaker is aware of is a certain function, or mapping, F such that $F(s_1') = s_1$. Suppose further that the hearer builds up some abstract situation s_2' as the result of the discourse. The situations s_1' and s_2' may be similar, or even the same, but it is not what the speaker intends to achieve. The speaker wants s_2' to have a similar, if not the same, significance to the hearer as s_1' has to himself. In other words he wants the hearer to discover F, not merely s_1. In fact, s_1 may not be observable by the hearer. The hearer may relate his model s_2' to some other part of reality s_2, but if he understood the message, there would be $F(s_2') = s_2$. In this sense what is communicated is the intension of F.

No matter, however, what is the ultimate intention of the speaker, the primary role of a discourse is to pass information. In this limited sense the task that the hearer is facing can be characterized as follows. Assume a certain set of characteristics of the speaker, taking into account his mental disposition, cultural background, social position, etc., etc., and his attitude toward the hearer. Then considering the present discourse situation context try to restore the speaker's discourse model from his message. The first step in this direction would be to grasp the message's literal meaning by building up a proper representation of this meaning at some SLA_n level. This is the aspect of the discourse analysis we want to discuss in the rest of this chapter.

6.2.1. Cohesion and coherence

When the speaker communicates about some situation, he usually needs a number of sentences to describe all relevant aspects of this situation. What connects his utterances is the underlying meaningful situation that he is discussing. That assumption guarantees the **coherence** of the discourse It has been observed (see, for example, Brown & Yule (1984)) that whenever people approach a spoken or written fragment of the language, they tend to consider it initially as a coherent discourse. There may be many reasons for doing so, but the simple physical proximity of utterances, or some specific typographical organization of a written paragraph are probably dominant. Brown and Yule (1984) show examples where some purposely incoherent paragraphs are nevertheless interpreted as discourses as long as the reader can make sense of them. These examples suggest that, unless forewarned, we try to interpret a connected paragraph as a coherent discourse as far as possible. In doing so we almost automatically build fragments of the representation of the content of the alleged discourse message. This image derives gradually, and each sentence adds some new elements to it. If initially there appear only some seemingly unrelated points, we tend to expect that the rest of the discourse provides the necessary perspective rather than abandon the whole discourse as meaningless. The latter reaction is however, ultimately unavoidable when one of the following takes place. Either the discourse ends abruptly without delivering enough data to build a coherent image of its contents or the narrative takes an unexpected turn awakening our attention that something is going awry

The notion of coherence appears then to be at the essence of any discourse. What we need is some concrete means for telling a discourse from a mere collection of unrelated language expressions For that purpose we have to simplify our original problem. First, we assume that the speaker does not intend deliberately to mislead us by producing incoherent "discourses". Second, we shall consider a discourse to be any connected fragment of language that cannot be proven incoherent using some presumed criteria which may include the two mentioned above. There is an important aspect of any discourse which may help us to build that criteria, and seems to be, at least theoretically, amenable to computer analysis. A (coherent) discourse exhibits internal connectivity between sentences: it is

cohesive. Although cohesion is a necessary feature of discourse. it unfortunately does not suffice for classifying a collection of sentences or utterances as a discourse. Even with the simplifying assumption above. the cohesive relations in text should be considered with great caution. We have no reason, however. to underestimate the role of cohesive links in discourse. Two general, and rather intuitive principles can be observed that offer some guidance in selecting proper cohesive relations in text. The principle of locality of context instructs the hearer to look for the narrowest possible context to find interpretation of some discourse utterance. In this sense, when presented with the following fragment

(5)   *The baby cried.*

      The mommy picked it up. (Brown & Yule 1984)

mostly without hesitation we resolve the pronominal reference of *it* as referring to *the baby* mentioned in the first sentence. The principle of analogy is slightly less specific. It instructs the hearer that unless told to the contrary. he should assume that things do not change unexpectedly from one sentence to another. This principle allows the hearer for using a partially developed representation of the discourse content to interpreting incoming utterances. Taken together. the two principles seem to promote the tendency of uncritical pursuit after a coherent sense of any piece of language. When none of the principles is applicable at some moment. we may say. without changing the representation of the discourse content constructed thus far. that something went wrong and what we are experiencing is not a coherent discourse. Unfortunately. things do not work so simply. Another aspect of discourse. the discourse topic. has a great influence upon deciding on discourse coherence. The following subsection summarizes briefly some of the most important aspects of this notion.

6.2.2. Discourse topic and speaker's topic

Informally. a discourse topic is what the discourse is about. We may talk about politics. cooking. Prolog programming. how to register for the next semester. and so on. Clearly. such characteristics can be more or less to the point. and the degree of precision strongly depends on our present interest

and expertise in the subject matter. Ideally, we would require a discourse topic to be a single compact expression (a proposition) that acts as a precise paraphrase of the discourse content. In this sense the topic may be thought of as a title of a discourse. Unfortunately, this approach is not very realistic in general. We often experience discourses for which a single phrase title, if one can be given at all, must be selected from some wider context, and thus cannot be as precise as one would wish it to be. Complications compound further when one considers a conversational discourse. Although we can often describe *what they are talking about*, it can only rarely be precisely expressed in just one phrase or sentence. That difficulty stems, in part, from the fact that each party in a discourse may pursue his own subject so that the resultant topic becomes some rather casual combination of these. But even in one-speaker discourse, the ultimate definition of what is being discussed may differ between the speaker and his audience, as well as between members of this audience. From the point of view of a natural language understanding system, it would be most desirable to identify each speaker's topic and then attempt to synthesize what is being discussed. To accomplish the first of these goals we shall once again, simplify our model and put the computer in the role of the sole hearer of a one speaker discourse. Now our task is limited to extracting the speaker's topic from his narrative.

We have already expressed that an attempt to provide a single phrase topic often makes no sense. Instead a notion of topic framework can be suggested † a notion of topic framework can be suggested as a tool to derive the discourse topic.† A topic framework will consist of these elements and aspects of the context that are activated or evoked in the discourse. It cannot be, however, just an arbitrary set of entities. In fact the elements of the topic framework should be only these that have the key significance to the subject matter, and they should additionally be organized to reflect the discourse internal structure, see Grosz and Sidner (1985). Pollanyi and Sch: (1984). Theoretically, however, any collection of active elements of context that include these key entities would suffice. Although the topic framework narrows our investigation space, it may be still a non-trivial task to

---

† This concept is recognized by numerous researchers, see Grosz (1977), Webber (1979), Grosz and Sidner (1985), though particular definitions may vary. The term of topic framework is used here after Brown and Yule (1984).

determine the topic from even the most accurate topic framework. Indeed the content of a discourse alone may not give us enough clues to what is being actually described. Consider, for example, the following paragraph

> The procedure is actually quite simple. First you arrange things into different groups. Of course, one pile may be sufficient depending on how much there is to do. If you have to go somewhere else due to lack of facilities that is the next step otherwise you are pretty well set. It is important not to overdo things. That is, it is better to do too few things at once than too many. In the short run this may not seem important but complications can easily arise. (Brown & Yule 1984)

Although we can accept this passage as a piece of a discourse. we feel uneasy when asked about its topic. One may seriously wonder whether the passage is coherent, or even connected. until told that it is about washing clothes. This example is a little artificial but it clearly shows how the discourse topic contributes to our understanding of language. Situations like this are not so rare in reality. When you come in to the middle of some conversation, after listening for a while. you may ask *wait a minute. what are you talking about?* demanding specification of the topic, so that you can participate in this discourse. Such a situation happens when the discourse model built by the hearer differs considerably from that of the speaker. A discourse model should therefore contain both the discourse topic and the topic framework. Actually it will contain all the entities evoked in the discourse (Webber 1979) either directly or from some wider context. From the hearer's perspective it may then look reasonable to maintain a topic framework as an approximation of the speaker's topic. When the hearer feels at loss with the conversation he may always interrupt. demanding some help. And although the topic framework does not yet guarantee finding the topic. it can. as we shall see. help in building the discourse model by selecting proper cohesive links between utterances

The question naturally arises how to indicate the members of the topic framework and select them from among all other elements of discourse model. We concentrate here only on these entities which are more or less directly evoked in the discourse message. When one considers a discourse message as a collection of propositions made by the speaker then the two following notions are of particular importance. the sentential topic. and the topic entity. or discourse theme. A sentence topic may be identified with the object that is referred to by the most exposed constituent of the utterance. In

written language the phrase occupying subject position is normally so considered This may get

altered however by *topicalization* (Radford 1983) which plays the role of shifting the usual sentence

stress pattern onto some selected phrase Thus in

(6) *That small green book. I gave it to you a few days ago*

the sentential topic is the phrase *that small green book* rather than the subject *I* The notion of

sentential topic is an important part of the transformation F_n It may help when selecting proper local

co-referential links between adjacent sentences of a discourse We shall say more to this point in the

next two sections.

The notion of sentential topic may be augmented over the whole discourse creating the concept

of the discourse main character - the topic entity There may be more than one topic entity in a

discourse but they tend to be topicalized or thematized in considerable fragments of discourse If such

a topic entity could be traced in a discourse, it would offer another criterion in resolving anaphoric

references in text.

Another discourse phenomenon which may significantly complicate the image we have created

thus far is topic shift The problem of discourse topic shifts also includes related phenomena of focus

ing and changing internal structure of discourse see for example Grosz (1977) Pollanyi and Scha

(1984). Grosz and Sidner (1985) In general the more precise specification of discourse topic the

more sensitive it is on further development of the discourse Usually a change in discourse topic will

affect the discourse theme a new topic entity may enter the discourse, or the present theme may get

moved into a new environment This in turn, can influence the ability of a natural language under

standing system to analyze the rest of discourse and a period immediately following the actual change

may be critical in that matter. These changes include not only the top-level topic shifts but also vary

ing levels of focusing created by elaborating on some aspect of the current focus recovering from

digressions flashbacks interruptions etc see Grosz and Sidner (1985). The problem of discovering

discourse topic boundaries is not an easy one and we shall mostly ignore the issue in this thesis Some

clues can be provided by the typographical organization of the discourse. and by the use of special tex-

tual markers indicating a change of the present subject matter. like *Another question.... Generally....*

Then . etc When these textual markers are unavailable. lack of cohesive ties to the part of discourse

that has already been processed may indicate a shift on the discourse topic. We shall return to this

latter question by the end of this chapter In the following section we discuss discourse cohesion prob-

lems which will play a central role in the transformation F_n

6.3. Cohesive relations in text

Although chapters 4 and 5 were almost entirely devoted to the problems of formal description of

some classes of cohesive links in discourse. there are many more sources of textual cohesion which we

did not mention and some of them will not belong to the domain of the transformation F_{n-1}. Some

more explicitly marked cohesive relations could be traced without difficulty by earlier transformations

F_i, $i < n-1$ Some others we would pass to the discourse level analysis of the transformation F_n We

present a list of textual cohesive relations compiled mainly from Webber (1979) and Brown and Yule

(1984) For each position in this list we provide a brief comment where. in our opinion. that particular

class of cohesive links belongs in the Stratified Model: to some early transformation F_i, $i < n-1$. to the

transformation F_{n-1} or to the transformation F_n

Textual cohesive relations can be divided into two general (but not exhaustive) classes of *expli-

cit markers* and *implicit ties* Within the class of explicit markers Brown and Yule (1984)

differentiate four smaller types. *additive markers* (and. or. furthermore. ...). *adversative markers*

(but. however. on the other hand.). *causal markers* (so. consequently. ...). and *temporal markers*

(then. finally.) Many of them can be detected and properly represented by some pre-F_{n-1} transfor-

mations. preferably at a syntax-dominated level. Others. like most of the casual and temporal markers.

after an early classification on a syntactic level. should undergo contextual evaluation with the

transformation F_{n-1} Temporal markers were not discussed in Chapter 4. but their treatment can be

formalized in a manner similar to the treatment of definite noun phrases There is not much more to say about this class of cohesive relations. except that they will serve as evidence of connectivity for the transformation F_n

The class of implicit ties is more interesting because these links, have to be handled by the transformations F_{n-1} and F_n This class includes subclasses of endophoric (in-text) relations. and exophoric ties connected to extra-textual objects. We shall mostly neglect the latter category in this chapter. although our reference scheme from Chapter 4 can handle both types (recall our definition of the context-setting sentence from the beginning of Chapter 4). The endophoric relations are what we generally ascribed to the domain of the transformation F_{n-1}. They include backward in-text references (anaphora). and forward in-text references (cataphora) The anaphora is the best developed type of implicit cohesive relation and we discussed numerous kinds of anaphoric links in Chapter 4. For a more comprehensive list the reader is referred to relevant literature. for example Webber (1979). In the classification assumed in this thesis. cohesive in-text relations considered by Webber (1979) fall into the larger class of direct. single-level references (cf Chapters 4 & 5 for appropriate definitions) As such they all can be at least theoretically handled by the transformation F_{n-1} in the form presented in Chapter 4

What remains for the transformation F_n in the area of text cohesion? The two general classes of text connectives we discussed above do not exhaust all types of cohesive links that can be found in discourse Indeed more sophisticated kinds of inter-sentential links can be found in natural language discourse and we signal these below together with some representative examples Whenever a *classification* comment is present. we refer to that level in the Stratified Model which is most likely to come in contact with the problem

Advanced cohesive relations in discourse

(a)   part-whole relation

John bought a new car yesterday   The engine works beautifully

classification: indirect reference: F_{n-1}.

(b)   collocability relation (e.g. Monday-Tuesday)

On Monday I went to see John. Tuesdays are reserved for golf.

classification: possibly handled by F_{n-1} as indirect reference, although the final decision must

belong to F_n

(c)   clausal substitution

John seeks a unicorn. So do I.

classification: F_{n-1}, not elaborated in this thesis.

(d)   comparison

All incoming applications will be considered. Preference is given to Canadian citizens

and permanent residents in Canada.

classification: a more complicated case of indirect reference.

(e)   syntactic repetition

We came in. They came in.   (Brown & Yule 1984)

classification detected early, selected by F_n

Besides the above relations, an observable consistency of tense or stylistic form may be regarded as

an indication of discourse continuity. This list, though incomplete, does not seem to leave much for

the transformation F_n. Observe, however, that in three cases: (a), (b), and (d), we suggested the

classification into the wider class of indirect references. We believe that computation of such "missing

links" is the most difficult problem in text cohesion. Recall from Chapter 4 that an element y of a sen-

tence $S_2(y)$ referred indirectly to an element x in a context-setting sentence $S_1(x)$ if there was a piece

of knowledge, revoked by the speaker/hearer from his knowledge base, that somehow related y to x.

This process normally involves an inference which, as in cases (a) and (b) above, can be standardized

to some degree. In general, however, detecting indirect references will require some non-trivial mental

effort, and in some cases the local context processing by F_{n-1} will not be adequate to properly capture

such links. Cohesive links that cannot be properly detected by the transformation F_{n-1} are classified into the category of *lost links* which belongs to the domain of the transformation F_n. We return to the problem of recovering lost links in discourse in section 6.5.

Except the category of lost links, and perhaps some other non-locally derivable connections, the transformation F_{n-1} is expected to deliver all possible cohesive relations found in a discourse onto the level SLA_{n-1}. What we consider the primary role of the transformation F_n is to select these actual cohesive links as intended by the speaker, and then build an unambiguous representation of discourse at the level SLA_n. We call this process **disambiguating in discourse**. The following section raises some issues that may be relevant in formulating this aspect of the transformation F_n.

6.4. Selecting proper cohesive links in discourse

There are many factors that influence the selection of cohesive links in discourse. We concentrate only on in-text anaphoric references, since they are the most natural continuation of the discussion which began in chapters 4 and 5. We do not propose actual solutions in the form of translation rules. Instead, we present and comment on a number of issues which will have a predominant influence on the ultimate formulation of the transformation F_n.

Suppose that the current discourse utterance $S_2(y_1, \ldots, y_k)$ contains k different anaphoric (or cataphoric) expressions $y_1, \ldots, y_k$, all of which have to be resolved in the previous (or incoming) discourse. Let X be the set of all objects brought into the discourse model accumulated thus far, and $A = \{x_1, \ldots, x_n\} \subseteq X$ be a non-empty subset of these such that each x_i is likely to be referenced in S_2.† We shall call the structure $S_1(x_1, \ldots, x_n)$ an **(active) context-setting sentence** for $S_2(y_1, \ldots, y_k)$. In the rest of this chapter we refer to $S_1(\cdots)$ plainly as a context-setting sentence, but the reader must remember that an active context may often appear only a subset of the entire

---

† The actual contents and structure of the set A will be determined in part by the structural organization of discourse model. We do not investigate these problems here. The reader is referred, however, to Pollanyi and Scha (1984), Cohen (1984).

available context as defined in Chapter 4. Observe that neither S_1 nor S_2 has to be a commonly understood syntactic unit. As we have pointed out in Chapter 4. S_1 accounts for the fragment of the discourse model accumulated up to the current utterance (for anaphora): and it may include new objects added in by the current utterance as well. In this way. we can include sentence internal anaphora in our framework. for example *John likes himself*. Assume further that for the S_2 to be properly understood. all of its k anaphoric expressions must find their antecedents among elements of A.

A word of warning is necessary here: elements of A are not to be identified with those "evoked" in discourse as suggested by Webber (1979). In particular. an individual. or a set of individuals. may not stand in one-to-one correspondence with a certain nominal construction used somewhere in discourse. The example below illustrates this point.

(7) a *Wash and core six cooking apples*

    b. *Put them into a fireproof dish.* (Brown & Yule 1984)

Although an object corresponding to the set of six cooking apples is present in the discourse model at some location l_1. at some later location l_2 this object gets replaced by a new object corresponding to the set of six washed and cored apples. The former object is no longer available as an antecedent. except when we explicitly refer to the location l_1. for example replacing (7b) above by (8).

(8) *John bought them in a supermarket.*

Indeed. the space and time conditions will be of some importance in selecting proper anaphoric links in discourse. For the moment. however. let us consider the reference situation consisting of the current utterance $S_2(y_1, \cdots, y_k)$ and the context $S_1(x_1, \cdots, x_n)$ where no guidance is given on how to coindex x_i's with y_j's. No constraints are imposed on the cardinality of either $R = \{y_1, \cdots, y_k\}$ or A. except that both should be finite. It is an open question whether $k \leq n$. as some of the y_j's may refer to the same x_i. As an example consider

---

Grosz and Sidner (1985).

(9)· a. *For Christmas. Bill bought Mary a cassette player and a blue dress*

  b. *She liked the small machine so much that she carried it with her everywhere.*

If (9a) is the sole context for (9b) then the set A consists of three elements: *Mary, a cassette player* and *a blue dress,* which are potential referents for the five element set $R = \left\{ \text{*she, the small machine,*} \right.$

she, it, her $\left. \right\}$ of anaphoric expressions in $S_2 = (9b)$. We have then $k > n$, and not all of the elements of

A are being addressed in (9b). In fact both *she* and *her* refer to *Mary,* while *the small machine* and

it take *the cassette player John bought* as the antecedent. The latter reference case is perhaps-not

that clear as one may say that *it* refers only indirectly to the cassette player with its immediate

antecedent being *the small machine.* Aside from this debatable case, we have to exclude clearly from

the above phenomenon, situations where the intermediate antecedent is all important. and a shared

ultimate referent appears contingent. Compare (10) below

(10) a. *John has passed the exam.*

  b. *He is proud of himself*

Here the reflexive *himself* refers to *he* in (10b), and only in a roundabout way to *John* by virtue of the

fact that *he* refers to *John*

Returning to S_1 and S_2, we can reason as follows. Let $S_1(A)$ be a context-setting sentence with

a non-empty, finite set A of objects likely to be referenced in upcoming discourse. Let $S_2(R)$ be the

fragment of the current utterance containing a finite collection R of anaphoric expressions to be

resolved against S_1. Then the number Θ of (theoretically) possible ways to tie S_2 to S_1 is

(11) $\Theta = |A|^{|R|}$

where $|\,|$ is the set cardinality function. The number Θ may be, to some degree, considered as a meas-

ure of connectivity (cohesion) between the current utterance and the rest of discourse. With $\Theta > 0$ the

cohesion will be considered as probable: this includes the case when $|R| = 0$. In such a case the sole

physical proximity of S_1 and S_2 is considered sufficient. However, when the context is empty, i.e.,

$|A| = 0$, no cohesion can be claimed. This is the situation where all of the elements of R, $|R| > 0$, refer somewhere outside A. If both A and R are empty, the state is undefined: no context is either provided or called upon, and the question of cohesion makes no sense there.

In example (9), excluding *Bill* from A as before, we have $\Theta = 3^5 = 243$ possibilities to interpret (9b) in the context of (9a). The observation that in an actual discourse situation (as seen by the hearer) no more than a few feasible connections are considered, and most often just one is assumed, is warranted by a number of constraints we often semi-consciously apply to rule out other interpretations. In (9), for example, we almost immediately exclude *she* and *her* as referring to anything but *Mary* on the sole basis of number/gender agreement between the pronouns and the name. To link *the small machine* to *the cassette player John bought* we have to fetch a piece of our general knowledge which says that cassette players are more likely to be described as machines than dresses and people. Finally, *it* can be ruled out as referring to *Mary* with some lexical/grammatical considerations. The choice between *the cassette player* and *the blue dress* as the antecedent for *it* in favor of the former is dictated by more sophisticated pragmatic issues, but the concept of recency: *a cassette player - the small machine - it*, and the particular structure of the sentence (9b) (topicalization) may suffice in many cases. However obvious a particular reference case happens to be, the transformation $\mathcal{F}_n$ must be equipped with an appropriate collection of rules to reduce the theoretically large scale ambiguity to just the number of possibilities an intelligent individual would seriously consider feasible. Below, we discuss a number of factors that will influence the formulation of such rules.

Generally we can divide these factors into three groups: factors of lexical/grammatical origin, factors which involve semantic/pragmatic considerations, and factors that derive from discourse coherence.

A: lexical/grammatical factors .

A straightforward number/gender agreement constraint plays a profoundly important role in this category. This constraint often significantly narrows the domain of pronominal anaphora In (12) below: the preferred antecedent of *she* is *Mary* on the basis of the "femaleness" of the latter.

(12) *John bought Mary a blue dress.*

   She looked very pleased.

Though simple and effective. the rule has its exceptions. especially when the personal/non-personal distinction is involved. Consider, for instance, the following:

(13) *The car has not got time to avoid the person, and he hits it.*

Here. the antecedent of *he* is *the car* while *it* refers to *the person* This is an example of a situation where the number/gender agreement rule has been overridden by some clearly pragmatic considerations.

   Another rule that we classified within the lexical/grammatical category comes from the Transformational Theory of Syntax (Radford 1983) The rule consists of three conditions: (A) Anaphor Binding Condition. (B) Pronominal Binding Condition. and (C) Lexical Binding Condition. According to these conditions the following pronominal anaphora are resolved as shown in (14) below (co-indexing indicates co-reference):

(14) *John$_1$ likes his$_1$ cat.* (by A)

   John$_1$ expects Fred to help him$_1$. (by B)

   He$_1$ does not believe John$_2$ could win. (by C)

   John$_1$ does not believe he$_1$ could win. (possible by ¬C)

The reader interested in formulation of these conditions in referred to Radford (1983) Other relevant works in linguistics regarding problems of definite anaphora. co-reference and logical form include that of Heim (1982) and Hornstein (1984).

B: semantic/pragmatic factors

This category is probably the most frequently employed to resolve local anaphoric references in discourse. Factors from this group are usually more decisive than those factors of purely lexical/syntactic origin, but they normally require a considerable amount of knowledge that is often not available from a discourse. We list five types of such factors which will account for a significant part of the transformation F_n. See also Webber (1979) and Brown and Yule (1984) for similar classifications.

The simplest rule in this category is derived from Fillmore's concept of *sentence frame* (Fillmore 1968). We shall call it here **the role matching constraint**. According to this rule, and the principle of analogy, an anaphoric expression should take as an antecedent that object in the discourse model which has been assigned a matching case in some recent utterance. Thus in (15) below, the anaphora of *he* is resolved to *John*, and not to *Fred*, as the former is responsible for the call, thus matching the role of "agent" with the antecedent of *he* who is responsible for the plea for help.

(15) *John called Fred because he needed help.* (Webber 1979)

The range of this phenomenon is not all clear. After all, with *he* referring to *Fred*, (15) may be still acceptable, although the first interpretation seems preferable. Observe that when the pronominal anaphora is replaced by a definite description the preference appears to be biased conversely.

(15a) *John called Fred because the man needed help.*

An advantage that we may note is the relative computational realizability of the role matching constraint. Supported by discourse coherence conditions this rule should find its place in the transformation F_n.

Another constraint that may be imposed on linking a current utterance with remaining discourse has already been considered in Chapter 4. Recall that a certain class of objects appearing in context-setting sentences was identified by two properties P and Q attached to them. In other words, an object **x** belonging to this class, and such that $S_1(\cdots x \cdots)$ is a context-setting sentence for some current utterance S_2, has been characterized as a unique individual such that $P(x)$ & $Q(x)$ held.

Suppose then that y is an anaphoric expression in S_2, i.e., $S_2(\cdots y \cdots)$, and that in the context-less (literal) translation of S_2 (see Chapter 4) it is identified with an object y such that $P_1(y)$ & $Q_1(y)$ is maintained. We have implicitly assumed in Chapter 4 that for y to be resolved to x, it is required that the properties P and P_1 are compatible. We call this condition **the Property P Matching Constraint.** This constraint instructs the hearer to link *the poor animal* with *the donkey, John owns* rather than with *John*, in (16) below.

(16) *John owns a donkey.*

  The poor animal has to work every day.

Here, the property P of being a donkey is compatible with that of being an animal. Without some indication to the contrary *John* is naturally classified as a human, and this is additionally stressed by his position in the ownership relation. One of the weaknesses of this otherwise powerful constraint is the difficulty of deciding actual compatibility between properties P and P_1. Compatibility decisions often require a considerable amount of knowledge and/or reasoning and may depend on the describing convention set up by the speaker in a particular piece of discourse. Compare (16) with (17) below.

(17) *The kid has tortured this poor kitten again*

  The beast must be kept off my garden.

Although in an ordinary discourse situation the hearer normally has no particular problems with computing the compatibility of P and P_1 as intended by the speaker, the rule may prove insufficient in many cases. What happens if more than one object in context have their properties P compatible with P_1? In such a case we shall look for some other factors which may appear decisive. Among these **the Property Q Matching Constraint** may offer some help. The Property P Matching fails to produce a unique link between S_1 and S_2 if either: (i) for an y such that $S_2(\cdots y \cdots)$ and $P_1(y)$, we have a set of objects $x_1, \cdots, x_k$ such that $S_1(x_1, \cdots, x_k)$, and for every $i = 1 \cdots k$ if $P_{x_i}(x_i)$ then P_i is compatible with P_{x_i}, or (ii) there is no such x_i in S_1, which indicates the hearer's inability to discover the speaker intention (if there is one at all). The latter case may be considered as a special case of the

former where all relevant objects in context are in focus, and the property P_1 is as useless in this respect as if it was not given at all. In other words the property P_1 is either too general (use of personal pronoun. etc.). or is intended to introduce a new. not indicated earlier. characteristics of some object in context. These are the conditions in which we may wish to examine properties Q of the objects involved. Compare. for example. (18) and (19).

(18) a. *John was busy all day with a paper. and it was Bill who prepared the supper that night.*

    b. *He did not know much about cooking.*

(19) a. *John was busy all day with a paper. and it was Bill who prepared the supper that night.*

    b. *He worked on his project for a Prolog course.*

In both cases, the context S_1 (set by (18a) and (19a) respectively) contains two relevant objects *John* and *Bill* with properties P of *maleness*. and properties Q of *being busy all day with a paper* and *preparing the supper* respectively. In (18b) the anaphoric pronoun *he* might be claimed at best to possess the same *maleness* as the property P_1. which cannot help us to select between *John* and *Bill* for the actual antecedent. But the object referenced by *he* has the property Q_1 described as *not knowing much about cooking*. which allows us to confidently pick up *Bill* as the antecedent. For similar reasons we link *he* with *John* in (19).

A number of variations to the above rules can be produced. for example crossing PQ or QP matchings. The reader should not have much of a problem deriving suitable examples. We believe however. that these cases can always be reduced to either P-matching or Q-matching. In (20b) below. the anaphora of *the guy* is resolved to *the cook John hired* on the basis of property Q matching. because the cook can be inferred to perform some cooking between the moment he was hired by John and the time the supper was ready.

(20) a. *John was busy all day with a paper so he hired a cook.*

    b. *The guy prepared a delightful supper.*

The notion of property matching may be still generalized to include some more subtle linking restriction, sometimes known as **selectional restrictions**. These will, almost invariably, involve reasoning in the hearer's knowledge base, most often outside the discourse model. As an example consider pairs of sentences (a-b), (a-c), and (a-d) in (21) with respect to anaphora *it*.

(21) a. *John stopped the car in front of his house.*

  b. *It badly needed a wash.* (it = the car)

  c. *It looked deserted.* (it = John's house)

  d. *It badly needed new paint.* (undecided)

There are two additional types of factors in the semantic/pragmatic class we would like to indicate here. They involve mostly pragmatic considerations and a good orientation in the discourse model accumulated thus far. One of these factors, sometimes called **the scene shift restriction** see for example, Grosz (1977), and Webber (1979), instructs the hearer to look for possible antecedents for an anaphoric expression within the limits of the current scene or situation described by the speaker's narrative. For example, *the door* in (22) below is preferably resolved to John's house door rather than to the car's door, as the effect of scene change between (22a) and (22b).

(22) a. *John left the car in front of his house*

  b. *He slowly opened the door.*

The other type of factor that we call **situational restrictions** complements the scene shifts in that it provides the hearer with some guidance on how to select anaphora antecedents within a scene. These factors are usually firmly set in the hearer's discourse model, although some external knowledge may be used.

(23) a. *John stopped the car in front of the garage*

  b. *He managed to open the door and got in.*

  c. *He managed to open the door and got out.*

With the context set by (23a) *the door* in (23b) will normally be referred to the garage door, while in

(23c) to the car door

C: discourse coherence factors

The factors influencing anaphora resolution as we presented them thus far made only a limited

use of the aspect which dominates the hearer's understanding of a discourse: the discourse coherence.

The fact that we can find locally optimal anaphoric links does not guarantee that our discourse model

is coherent, much less that it accurately restores the speaker's intention. In fact, all hearer's decisions

are subordinated to what the hearer believes to be, at present, the coherent drive of a discourse. This

attitude may change more than once as the discourse progresses, which may frequently involve a par-

tial reformatting of the discourse model. The process of rebuilding a discourse model is often a costly

endeavour that may potentially cause serious perturbations in the hearer's ability to understand the

discourse fully. Although it is possible for the hearer to maintain more than one competing version of

discourse model at a time, he ultimately selects one model and further builds upon it. Therefore, the

hearer will first attempt to fit the incoming discourse into his model before he seriously considers any

changes in the model, and then he tries to minimize the extent of these changes.

Most of the informal rules for anaphora resolution that we have discussed in this section were

ultimately based on hearer's preferences and thus bore a potential for misinterpreting the speaker's

intention. Discourse coherence factors also are relative to the hearer's understanding of the discourse,

but these factors address the global structure of the hearer's discourse model. We classify these fac-

tors under a general criterion called **the Consistency of Interpretation Constraint**. The criterion has

a universal character, and will result in a number of specific rules of the transformation F_n. We explain

its role by example

(24) a *John is looking for Smith's murderer*

    b *I wonder where he is now*

As one may expect. the problem lies in finding the proper antecedent for the pronominal *he* in (24b)
Assuming that the relevant context includes only *John* and *Smith's murderer* we have to decide
which of them is actually referred to by *he*. Previously discussed criteria can give us one or the other
but considering the overall discourse coherence only one of them. *John*. appears the likely candidate
Suppose. therefore. that from the discourse preceding (24) as well as from (24a) the hearer draws the
conclusion that *Smith's murderer* has no actual extension in reality. and that the description is used
non-referentially by the speaker. In the terminology of Chapter 4. this means that (24a) assumes its
imperfect interpretation. Now suppose that the pronominal *he* in (24b) is taken as referring anaphori-
cally to *Smith's murderer*. Since (24b) has no clear imperfect reading. and *he* stands in a referential
position. in linking the pronominal to *Smith's murderer* we would break the consistency of the
hearer's discourse model by making the Smith's murderer an explicit. though unknown. individual
That is why we would rather relate *he* to *John*. especially because we do not have to sacrifice our
discourse model

There is another point to raise from the discussion above In (24) we could escape revision of the
discourse model because we perceived an alternative continuation of discourse What happens. how
ever. when all alternatives consistent with the current model are excluded? In other words. let M be a
coherent discourse model maintained by the hearer and let γ be the utterance in question Let further
$\Gamma = \{\gamma_1 \ldots \gamma_t\}$ be the set of all possible interpretations of γ relevant to the present discourse † If

for every $i = 1 \ldots t$ the model $M \cup \{\gamma_i\}$ is no longer coherent then we have to revise M to some M'

such that both M' and $M' \cup \{\gamma_i\}$ for some $1 \leqslant i \leqslant t$ are coherent. This process that we call **the**

Backward Adjustment Process may have varying impact on the structure of M. but it would be a
simplicistic view in general to expect only some local changes to the discourse model. The process will
normally involve a retreat from some previously made decisions up to and including the decision

---

† This relevance should conform to the discourse topic discussed earlier in this chapter.

directly responsible for the present inconsistency. In (24). if we decided that Smith's murderer

after all. used referentially. we would have to revise the discourse model. beginning from the instant

where. for the first time. we assumed the non-referential status of this concept. On the other hand.

when there is a $\gamma_k \in \Gamma$ such that $M \cup \{\gamma_k\}$ is coherent then we extend M by γ_k by computing

$M \leftarrow M \cup \{\gamma_k\}$. and relate the forthcoming discourse to the new M. Observe that from the computa-

tional viewpoint we have to keep track of all such extensions to M as they are potentially of interest to

the Backward Adjustment Process.

When none of the γ_i's fits into M the hearer has still at least three options for continuation. He

may ignore γ assuming that the speaker misformulated his intention. In this case no immediate revi-

sion to M is necessary. but we may endanger our understanding of the speaker's message if we. not

him. were in error. Secondly. the hearer can create a revised model M' while maintaining the old ver-

sion. especially when he is not sure who is at fault. and wait for clarification in upcoming discourse

This is the safe option. but the cost of maintaining concurrent discourse models can be considerable.

Finally the hearer can forget M concentrating entirely on the new model. All these approaches have

their advantages and disadvantages and we postpone the ultimate decision for future research.

Nevertheless. the Backward Adjustment Process will create an important part of the transformation

F_n

With this we end our discussion of the third. and the last. of the classes of factors influencing

resolution of anaphoric ties in discourse. Before we close the present section. however. we shall briefly

mention of two additional general principles that are of some significance for the transformation F_n.

Two fairly general. but sometimes mutually contradictory. principles have been observed in con-

nection with the problem of discourse anaphora. see Webber (1979). or Brown and Yule (1984). The

first. called **the principle of nearest entity** maintains that in a search for an anaphora antecedent the

preference should be given to that object from the set A of possible antecedents which has been most

recently mentioned in discourse. According to this rule. the anaphoric reference of *they* in the following

example is to be linked to *the survivors* rather than to *the rescue team* on the sole basis of the

recency of the former.

(25) a. *On the fifth day the rescue team found the survivors.*

    b. *They were very exhausted.*

A somewhat contradictory view is enunciated by the other principle. sometimes called **the principle of topic entity**. that instructs the hearer to instantiate an anaphora at the current discourse topic entity (see section 6.2). In (25a) the phrase *the rescue team* is thematized which makes it. lacking evidence to the contrary. the topic entity of the discourse. Application of the latter principle to (25) will have then the completely different effect on resolving the reference of *they*

As one may expect these two principles often do not, have a decisive impact on anaphora referent taking process They can be regarded as "last sort" instances when the more discriminating factors can no longer be called upon The principle of topic entity should be used only in situations where a topic entity is clearly perceptible in discourse. and in this sense it has priority over the princi ple of nearest entity

Selecting proper cohesive links in discourse and modelling the discourse representation accord ingly will be the main task of the transformation F_n But this does not seem to exhaust the problem of arriving at the ultimate discourse representation at the SLA_n level There remains a problem of recov ering **lost links** in discourse lost in the sense that the transformation F_{n-1} cannot deliver them We discuss this problem briefly in the following section

6.5. Recovering lost links

Most of the examples of "missing" cohesive links in discourse given in linguistic literature (see. for example. Webber (1979). or Brown and Yule (1984)) can be classified as instances of what we defined in Chapter 4 as indirect reference. The problem of indirect references can be characterized as

follows. Let $S_1(x)$ be. as before. a relevant fragment of discourse model M. considered as a context-setting sentence Let $S_2(y)$ be the current utterance with an anaphoric expression y. Assume. for sim plicity. that x is the only object in context available for anaphoric reference. If there is no direct correspondence between y and x. the hearer may nevertheless examine his knowledge base KB. $M \subseteq KB$. to see whether there is an object x' related to x. and such that x' can be directly referenced by y In other words the hearer has to augment his discourse model M to the effect that the relevant context for $S_2(y)$ becomes some $S_1'(x. x')$. and there is a direct link between y and x'. If such a process is possible we talk of an indirect reference between y and x As one may expect the hearer's ability to compute indirect references will greatly depend on content of his knowledge base. When the speaker produces indirect references in discourse he normally assumes that the hearer will be able to decode them In other words. he expects that the hearer shares appropriate part of his knowledge base.

The following is a typical example of an indirect reference.

(26) a *It was dark and stormy the night the millionaire was murdered.*

   b. *The killer left no clues for the police to trace* (Brown & Yule 1984)

The piece of knowledge sufficient to link the sentences (26a) and (26b) can be paraphrased as *murdering involves a killer* 'Some other connections like *murdering causes the police investigation* are also possible. What is really significant here is the fact that we can resolve such links *locally*. that is having only $S_1(x)$. $S_2(y)$. and KB we can always relate y to x This particular property of indirect references made them amenable to processing at the stage SLA_{n-1}.

There is. however. a class of missing links. we call these lost links. that cannot be resolved locally. This happens when the hearer knowledge base is not sufficient to relate x to y at the time the utterance S_2 comes in. The hearer may consider S_2 as the continuation of discourse because the connectivity factor Θ is positive. Besides. there may be some other links available between S_2 and S_1. Suppose that at some later point in the discourse. the speaker delivers that missing piece of informa-

tion so that the hearer can recover the previously lost reference. Consider the following example.

(27) a. *John returned to his house and turned the T V set on*.

    b. *The movie was really bad.* (the movie = TV-movie. or the movie = ?)

    c. *He decided he would never go to that theatre again* (the movie = theatre movie) .

The example is slightly simplified as the sentence (c) does not have to immediately follow the pair (a b). Also. the antecedent for the lost anaphoric link (here *that theatre)* may be explicitly available in S_1. but the hearer's knowledge base does not suffice for. or even precludes. the connection. To explain this condition let us make use of our schematic notation. Suppose the context contains two objects x_1 and x_2. i.e. $S_1(x_1, x_2)$ is the context-setting sentence. and y is an anaphoric expression in S_2. Assume now that the hearer's knowledge base KB relates y to x_2 (on the basis of property P or property Q matching. for instance). while the speaker's intention is to relate y to x_1. If the latter relationship is not supported in KB. it would be quite unjustified to insist that the transformation F_{n-1} should record that possibility too. We only may expect from F_{n-1}. except linking y to x_2. to create an alternative discourse model in which y is left pending (linked forward) When new information enters the hearer's discourse model or his knowledge base in general. he may try to reevaluate pending links. and restore lost references. Observe that the process of recovering lost links may involve revisions to discourse model. If the hearer has left some room in his model for such a link (because for example he expected it) then nothing significant would happen. But if he connected the discourse somehow differently than intended by the speaker. the new information providing the lost link (which the hearer did not perceive at all) may shatter the hearer's discourse model. This situation can happen at any time in discourse. and is by no means limited to anaphoric connectives. but we shall not investigate the matter here. Because of the non-locality of processing lost links. we think the process should belong to the transformation F_n. Note also that we have already gathered enough evidence to support the hypothesis that. in a practical natural language understanding system. the transformation F_n should indeed parallel at least that of F_{n-1}.

6.6. Conclusion

In this chapter we have highlighted a number of issues that we have found to be of particular importance for shaping the transformation F_n and the discourse representation level SLA_n. Not all relevant problems have been represented in our discussion, among them the organizational aspects of discourse as suggested by Grosz and Sidner (1985). More research is necessary in the area of discourse understanding, and this will create some direction for future work. One question seems to be in order. Is the transformation F_n the final transformation on the Left Side in the Stratified Model, and the level SLA_n the component of the ultimate stratum, $<SLA_n, UA_n>$? A definite answer to this question is not an easy one, but we think it should be positive. The reader may have already observed that the discourse representation on the SLA_n level, as we have tried to forward it in this chapter, will resemble in many aspects the concept of abstract situation (Barwise & Perry 1983). It will not be, in general, a single situation, however, nor even a tree-like structure of situations as advocated in the Theory of Situations. Rather, we would see the discourse representation as a collection of such situational structures, roughly corresponding to topic shifts in discourse. It should be also relatively clear by now that the discourse representation will not merely consist of skillfully coded speaker's message content. The representation must also mirror appropriate fragments of the hearer's knowledge base activated in discourse, for example that used to discover and select proper cohesive links. This amounts to the realization that the final discourse representation at SLA_n level will be directly derivable from the hearer's discourse model maintained at the end of discourse. As we have mentioned earlier, it is not compulsory to require a single representation be derivable for any particular discourse. If more than one concurrent representations are obtained, we should accept them as the evidence that the discourse in question has more than one possible readings. These will most often appear at different dimensions, for example literal and metaphorical, etc.

If we accept the above, another question prompts quite naturally. What would be the other part of so established ultimate stratum, i.e. what could we find in UA_n? We think it is too early to take any definite stand in this matter. It may seem not so unreasonable to consider UA_n as a system of

"real" situations. and thus regard the mapping M as Situation Semantics. Despite all the problems discussed in section 2.2. the situational approach may have some advantages over the possible world semantics in that respect. But, as we have said. we did not gather enough evidence to favor any of these theories. We feel that none of them, in its pure form as discussed in Chapter 2. is appropriate. The revisions we expect as inevitable may in fact give the beginning for a new semantic theory. the alternative we called upon at times. This is. however. a quite separate problem. outside the scope of this thesis.

Chapter 7

Addressing Computational Issues

7.1. Computer Realization of the Stratified Model

In our discussion we have stressed the computational perspective of our theory of meaning representation. The careful reader may have already assimilated some ideas on how to develop a computer implementation of various aspects of the Stratified Model. This chapter summarizes computational characteristics of the Stratified Model. and points out other related problems not addressed directly in the thesis. We do not propose an actual implementation for the Stratified Model since it would constitute a separate research challenge. Instead. we highlight a number of issues that have to be investigated in detail before any computer realization of the theory can be attempted. We discuss the computational side of the Stratified Model in general and then concentrate on the two most discussed transformations: F_{n-1} and F_n.

To implement the Stratified Model one must consider translating the source language from a natural form available at level SL into the ultimate representation at level SLA_n. and appropriate mappings of the "original" universe U onto various models including the ultimate model at level UA_n. Finally. the mapping M specifying the formal semantics of the language at level SLA_n should connect the Left Side of the Model with its Right Side. thus completing the implementation. Although each stratum $<SLA_i . UA_i>$. $0 \leqslant i \leqslant n$. has a mapping M_i such that $M_i : SLA_i \rightarrow UA_i$. which guarantees the meaningfulness of the Model. only $M_n = M$ must actually be made explicit in a computer implementation. This observation has a central significance in assessing the computational tractability of the Stratified Model. We have already pointed that the problems with contemporary natural language

understanding systems can often be identified with attempts to implement either an M_i such that $i < n$. or even to relate some two levels SLA_i (usually $i < n$) and UA_j (also $j < n$) that do not belong to the same stratum. i.e.. $i \neq j$. In fact. implementation of the Left and Right Sides should be coordinated as we suggested in Chapter 3. For every SLA_i to be created on the Left Side we should provide an appropriate UA_i on the Right Side so that the mapping M_i (however implicit) preserves the original semantics M_0 between SL and U. It is not necessary that the entire Right Side be explicitly constructed. With the exception of the ultimate level UA_n. and the universe U which is taken for granted. the remaining levels UA_i. $1 \leqslant i < n$. are mostly transparent for a system based on the Stratified Model. This becomes possible due to the implicit character of appropriate mappings M_i's We must provide, however. a method for decoding the universe representation at UA_n back into U. and the transformations G_n. G_1 on the Right Side must be given explicit instantiations (refer to Chapter 3 for more detailed discussion). As far as our major concern is to derive a meaning representation for a discourse. however. the Right Side may remain unspecified. except for the final model UA_n upon which we base our ultimate semantic mapping.

Suppose that we have already developed the Stratified Model with all necessary strata $< SLA_i$. $UA_i >$. $0 \leqslant i \leqslant n$. and described all required transformations F_i and G_i. as well as the mapping M In Chapter 3 we indicated that the Model is somehow possessed by an intelligent individual who. having also at his disposal an individual knowledge base KB. uses the Model for on-line processing of information entering either by means of source language discourse or by sensing the universe. or from the knowledge base. The flow of information may occur either in direction from SL to U (understanding). or backwards (communicating). often affecting the knowledge base and altering its contents. Let us concentrate here. as in preceding chapters. on the Left Side of the Model in forward processing. that is. from SL to SLA_n. The requirement of on-line processing restricts implementation of the transformations F_i's. We cannot afford a purely sequential scheme of the Left Side. It would be unrealistic to expect from a transformation F_i to wait until a full representation of a discourse is gathered at the level SLA_{i-1} before the transformation onto the next stratum could be attempted. In fact.

except for only some very early transformations $(F_1, F_2, \cdots)$, beginning from some level SLA_{s-1} all further transformations should work in tandem producing pieces of the ultimate discourse representation at the level SLA_n. This is not to say that every word, phrase, or even sentence entering SL should trigger a cascaded processing on the Left Side. The process could be rather compared to the counter mechanism where the move of a next digit ring occurs only after the preceding ring moves a number of times. For example, it would often be advisable to postpone the syntactic analysis of the input until a sufficient number of lexical items of a current phrase or sentence is available thus reducing the degree of non-determinism in parsing, see Marcus (1980).

This point is especially important with the categorial grammar CAT which assumes that all basic expressions (words) in a sentence have been assigned to syntactic categories. Thus the following sentence (pointed to me by Ray Jennings)

(1)   *John interviewed a treat.*

will result in two different parse trees in CAT depending on whether *treat* is assigned to the category t//e (common nouns) or IAV (IV modifying adverbs). Observe that in the latter case the verb *interview* appears intransitive. The transformation preceding F_{n-2} in the Stratified Model is therefore expected to deliver all alternative lexical/morphological analyses of the sentence in question by marking words with the syntactic categories they belong to. The syntactic level evaluation will determine which of these possibilities result in well-formed p-markers. As it may be expected, some of the early morphological ambiguities will persist beyond the syntactic level, into the discourse analysis. Consider, for example, a continuation of the discourse begun with (1) (R. Jennings, personal communication):

(2)   *John interviewed a treat.*

   But he was a poor judge of character.

Similarly, to the cooperation on morphological and syntactic levels, an extension to the discourse model by the transformation F_n may be considered only after the relationship of the current utterances to the preceding discourse has been determined by the transformation F_{n-1}. If we consider the

discourse (3) below (discussed in Chapter 4)

(3) (3.1) *John interviewed a man.*

    (3.2) *The bastard killed him.*

then we already know that F_{n-1} can produce at least two distinct discourse prototypes involving ana-phora *the bastard* and *him*. It is then up to F_n to decide which of these can survive when a wider discourse context is called for.

The process will not. however. be as regular as a counter. Some early transformations. including phonological. lexical. or even syntactic processing. which operate on fairly well-defined units (phonemes. lexemes. phrases). will be best realized sequentially. although this view may not be practi-cal. in general. We must always remember of the presence of the individual knowledge base which. by generating various presuppositions. may occasionally activate further transformations that build a higher level representation of some discourse fragment before that fragment is even fully sensed at SL. This happens when we recognize the beginning of some pattern language construction (idiom. com-monly used combination of words. etc.). and the knowledge base generates an expected continuation. In this sense the syntactic well-formedness of utterances. for example is not crucial for understanding them. If the actual continuation of discourse can be fit in our expectations further processing reduces to a mere verification. otherwise a revision has to be performed. if possible We discuss the problem of cooperating transformations in the next two sections where we examine computability of transforma-tions F_{n-1} and F_n

We know very little about the actual form transformations F_1 to F_{n-2} may eventually take. except perhaps that transformation F_{n-2} is identified with the categorial grammar CAT. We assume that transformations F_1 to F_{n-3} are computationally tractable (at least in theory). which assumption is quite reasonable considering localized effects these transformations bring about the discourse represen-tation. Lexical/morphological stage. as outlined above. would require an exhaustive lexicon on the domain of discourse in which each word is given all possible morphological interpretations. Values of

other morphological indicators such as number, gender or case should also be included, if applicable.

For example, the word *lot* may be expected to have at least the following entry in the lexicon.

(4)

        [lot:
                1. *category* t//e
                2. *category* IAV
                3. *category* TV]

For a detailed discussion on how to design and manipulate large lexicons the reader is referred to Cercone et al. (1983)

Implementation of CAT should not be a great challenge either,[†] but we do not exclude some other, and more powerful grammatical system employed as F_{n-2}. Although we believe that the problems of language processing one may expect to be covered by the pre-F_{n-1} transformations are more or less worked out already, we shall investigate further their nature and role in the Stratified Model. Let us therefore concentrate our attention on the transformations, i.e., F_{n-1} and F_n.

7.2. Computability of the transformation F_{n-1}

There are three aspects of the transformation F_{n-1} that merit special attention from a computational viewpoint. These aspects include: implementation of the rules which create the core of the transformation, (presently Rule 1 to Rule 12), automating the process of deriving context-setting sentences according to the recursive scheme given in section 4.10, and providing access to, and cooperation with, an individual knowledge base. To implement the reference rules one would generally need to solve the problem of λ-reductions where their computational tractability is questionable. In practice, however, a relatively simple subset of λ-calculus should suffice for this purpose. To the extent that we are concerned with the problem, implementation of λ-reductions has already been addressed in the computer literature, and concrete realizations have been suggested in standard programming languages

---

† Categorial grammars create a subclass of context-free grammars. The reader may note, however, that Montague's syn-

including Pascal and LISP (Georgeff 1984), and Prolog (Dahl 1984) The relative computational tracta

bility of the language Λ is favorably compared to the problems one has to face when attempting Λ

reductions in Montague's IL. See also Warren (1985) and further references therein

But even an effective algorithm for evaluating the rules will not have much significance until we

provide a method for computing context-setting sentences on which the rules operate. The recursive

formula of Chapter 4 is the first step in this direction. We have already shown the algorithm is well

founded and that it terminates. We believe it will prove programmable, especially when close coopera-

tion from the transformation F_n is assumed. What transformation F_n should provide is a considerable

reduction in the combinatorial explosion of possibilities generated by the formula. One of the most

important factors in this reduction comes from limiting the available context to that part which is

actually referenced by the current utterance S_2 (cf. the definition of context-setting sentence presented

in Chapter 6). An implementation of the forward context (the D_i' component of discourse prototype

cross-section) may pose a serious problem here. This part of the context derives mainly from presup-

positions made from an individual knowledge base.

Accessing and manipulating a knowledge base creates a separate programming problem. In addi-

tion to the well understood organizational questions of maintaining a large data base, one must solve

the general problem of inferences within the knowledge base. It would be impractical and unrealistic in

general, to expect the owner of a knowledge base to be always aware (read can access directly) of any

information that could be derived from what he knows or believes. However, the actual contents of the

knowledge base, as well as its internal organization, should guarantee that, in a specific environment

appropriate facts will surface as the result of proper inferences. This requirement, in turn, raises the

question of the quality and quantity of such inferences, i.e., how fine and in-depth they have to be in a

particular situation. Inferences which are too fine can be as harmful as inferences which are too shal-

low because they can undermine our ability to understand discourse at the level intended by the

speaker. Take, for example, the Schank-style (Schank 1972) conceptual information processing

tax is not context-free, see Partee (1976).

paradigm and the difficulties faced by the natural language understanding systems based on this approach (Schank 1975) In short the conceptual approach requires an advance specification of the entire world model on which language communicates as well as the level of detail at which inferences within this model are performed Although the framework can be quite successful in simulating some well-defined "toy-worlds" it proves hopelessly impractical in real world situations

Another problem directly connected with the transformation F_{n-1} is an implementation of the remote reference rules 11 and 12. What we require can be summarized as the ability of a natural language understanding system to distinguish between "ordinary" singular objects and superobjects. find a relationship between any two objects about to be linked (same level reference. subcontext reference supercontext reference unrelated. etc.) and eventually select proper coordinates. The above may be considered as an incremental process Initially. the available universe of objects is divided into levels with the relation of relative singularity defined in Chapter 5 The number and types of the coordinates involved will depend upon the discourse domain we are dealing with but one may expect that a number of different time and/or space coordinates will be essential in almost any universe This initial model may subsequently get extended by new objects and coordinates which situation may in turn require re-evaluation of the level structure For finite universes this procedure should not pose greater problems with implementation Observe that with the possible world approach we would need an undetermined number of indices even for finite universes

A successful implementation of F_{n-1} will depend however. upon considering all relevant aspects of some present discourse situation. in particular the discourse model This kind of information is normally not available for the transformation F_{n-1}. and may be visible only from the level SLA_n. The following section attempts to shed some light on the computational tractability of the transformation F_n.

7.3. Computability of the transformation F_n

The major task of this transformation is to build, maintain, and evaluate a discourse model that will eventually result in the final discourse representation at level SLA_n. Theoretically, this process amounts to the following: (1) selecting the "best" discourse prototype from among those generated at level SLA_{n-1} by employing various criteria described in Chapter 6. (2) uncovering remaining "lost links" within the selected prototype (or prototypes); and (3) revising so obtained discourse model (or models) by computing backward adjustments and indicating discourse topic boundaries. In practice, however, F_n should parallel F_{n-1}, and the process of constructing discourse model must be organized in an on-line fashion (cf our discussion in Chapter 3). Let M be the part of discourse model maintained by F_n at some instant c, roughly corresponding to the discourse prototype cross-section D_c (cf section 4.10). If M is unique, i.e., no alternative discourse models are considered at the time, the next step of the transformation F_{n-1} will produce a number of possible continuations M^j_{c+1} for M, $j \geqslant 1$, such that $M^j_{c+1} \in M \bigcup F_{n-1}(\delta_{c+1})$. F_n assists F_{n-1} in computing these continuations and then selects one (or more) that is believed to be intended by the speaker. Subsequently F_n will replace the old model M by a new version M' which derives from this selected M^j_{c+1}. The revision may include complicated adjustments to the old model, and often will not be just a simple extension with the meaning representation of the current utterance (refer to Chapter 6 for detailed discussion). Although F_n should continually assist F_{n-1}, it would not be efficient to replace the discourse model after every single step of the transformation F_{n-1}. Instead the transformation F_n should use a kind of "wait and see" strategy to intervene into the performance of F_{n-1} only if either (1) the current step of F_{n-1} resulted in too large a number of possible continuations of discourse so that it may soon escape control, or (2) F_{n-1} is sufficiently ahead of F_n so that the latter acquires enough confidence to select the best continuation for M. This technique, as we expect, will significantly reduce the necessity for backward re-evaluating of the discourse model. A closer determination of the pace at which F_n should augment the discourse model is not an easy problem. Partly, the decision belongs to a particular implementation, but for now, we know too little about the form the transformation F_n will eventually

assume

The limited "wait-and-see" strategy will not. however. protect the discourse model from further revisions. The transformation F_n must be equipped with a good backtracking mechanism to allow the process of backward adjustment of the discourse model (see Chapter 6) to be performed as smoothly as possible

A successful implementation of F_n will depend on the transformation's ability to recognize the proper continuation of discourse early. and to prune those options which are less likely. This capability includes an efficient implementation of the text cohesion factors discussed at length in Chapter 6. Most of these factors require a close cooperation with an individual knowledge base and its inference mechanism. To say something more definitive about the computational aspects of translation rules based on these factors we would have to investigate the organization and content of a knowledge base which is outside the scope of this thesis.

It would also be difficult at present to design an automated process for deriving discourse topic. tracing topic shifts. and detecting topic boundaries. The best we can hope for is to provide the transformation F_n with the means to compute topic framework (or focus) and distinguish so-called topic entities: see Sidner (1979). Grosz (1981). Grosz and Sidner (1985). The discourse topic may be generated then from the knowledge base by the means of .proper inferences. The success of this process will depend on both the content of the knowledge base and the discourse model maintained at the time. /

Finally. the transformation F_n should continually compute the active context. i.e.. that part of the discourse model which is most likely to be referenced in an upcoming discourse. An active context consists. in part. of a representation of some recent fragment of discourse. and of topic framework. Maintaining such a notion will greatly reduce the ambiguity of the transformation F_{n-1}. and subsequently speed up processing on the Left Side of the Stratified Model. More concrete decisions on the actual size. content. and organization of topic framework and active context should be left to a particu-

lar realization of the Model

7.4. Conclusion

By and large, we are quite optimistic about the computational prospects of the Stratified Model The Model overcomes some of the difficulties of Montague's approach in providing a discrete and more comprehensive scheme of language processing. Two aspects of our approach are especially noteworthy as seen from the computational perspective. The intensional logic (IL) has been replaced with the language Λ which promises computational tractability. This is in part due to the corresponding universe model at UA_{n-1} from which the language takes its interpretation. The highly problematic concept of intension as function over possible worlds is challenged with the local, discrete notions of a level coordinate and relation of relative singularity. Whether Λ will ultimately prove to be no more difficult to deal with than first order logic remains to be seen. To fully answer this question one has to define a semantics for Λ, in particular for Λ's second order operators like **imp** and **att** This issue is left for the future research. The Model maintains the flavor of a formal theory, in contrast to the mostly intuitive character of Theory of Situations. Of course much work remains to be done including development of a semantics for the discourse representation at level SLA_n and a closer investigation of the Right Side of the Model. All of this will be discussed in the following chapter where we try to define future directions for this research. A fuller implementation of the Stratified Model must be likewise postponed

Chapter 8

Future Directions and Conclusions

8.1. Future Directions

We examine four directions in which the work reported in this thesis may extend. Further investigations into the problems of inter-sentential dependencies that belong to the domain of the transformation F_{n-1} is one direction. Related to these investigations, one may wish to examine the problems of text cohesion discussed in Chapter 6 and prepare a formal description of the appropriate fragment of the transformation F_n, preferably in the form of translation rules similar to those defined for the transformation F_{n-1}. This effort would constitute another research direction which should also address problems of discourse organization, discourse topic, discourse theme and modelling discourse representation. A third future investigation would be the further development of the Stratified Model including specification of remaining transformations F_i, $i < n-2$, G_j, $j=1 \ldots n$, and M with special emphasis on the Right Side of the Model, i.e. that devoted to modelling the language denotational base. Finally an efficient implementation of the theory may be attempted.

In the rest of this section we list a number of discussion points that should be considered for these extensions. Many of them have already been indicated at various places in this thesis.

8.1.1. Computing Cohesive Links in Discourse

In the concluding section of Chapter 4 we acknowledged that Rules 2 to 10 account only for a subset of possible reference situations that can be defined between parts of discourse. In particular,

rules for dealing with *one*-anaphora; clausal *it*, and elliptical constructions have yet to be constructed and related to the five context situations distinguished there. A more complete list of types of discourse anaphora can be found in Webber (1979). This line of research offers perhaps the best established direction for future work.

A related problem is to how formally describe new types of context situations, and thus prepare the ground for more rules within the transformation F_{n-1}. In Chapter 4 we have already distinguished perfect contexts, imperfect contexts, attitude report contexts, contexts with proper names, and conditional contexts. Although we do not report it in this thesis, we have been investigating another context situation which we tentatively call the *creative context*. Roughly, a creative context may arise in connection with *creative verbs*, like *imagine, paint, write* etc., which can produce non-referential readings of utterances containing them with respect to the object position, as in *John has painted a unicorn* (recall also our earlier discussion of the verb *imagine* in Chapter 4) The notion of creative context requires more research before appropriate translation rules could be formulated. Nonetheless, it establishes another exciting line of development for the transformation F_{n-1}.

There are a few additional aspects of the transformation F_{n-1} that need future investigation. In the area of indirect references (section 4.8) we may expect some formal guidance as to when an indirect reference could be anticipated between two terms. In forward references (section 4.9) a further specification of the context D'_i will be required. Similar consideration must be given to the general processing scheme suggested for F_{n-1} in section 4.10

A truly challenging research direction emerges from attempts to extend the present version of the Theory of Names and Descriptions (TND) discussed in Chapter 5 beyond strictly nominal terms. It would be interesting to see how the concepts of superobject and coordinate could represent objects denoted by expressions from categories t/e, $t/(t/e)$ and others. Beside a number of new remote reference rules, the research may result in a new approach to natural language semantics that, in turn, may give the beginning to the formal definition of the mapping M on the ultimate stratum $<SLA_n, UA_n>$

This includes formal extension to the meaning representation language Λ at the level SLA_{n-1}.

8.1.2. Selecting Proper Cohesive Links in Discourse

The job of selecting proper cohesive links in discourse belongs to the transformation F_n that (formally) follows F_{n-1} in the Stratified Model. Ideally. at a particular reference situation $<S_1. S_2>$ with the context S_1 and the "current" utterance S_2. given a number of alternative connections between S_1 and S_2 provided by F_{n-1} in the form of discourse prototypes. the transformation F_n should select that continuation which an intelligent addressee would assume as intended by the speaker. In Chapter 6 we discussed a number of factors that influence in this matter. The most obvious direction for future research would be to formalize these factors into a set of well-defined translation rules. We expect these rules would assume the conditional form of *IF* <condition> *THEN* <decision>. The <condition> part is all important. Its primary role will be to restrict a rule's applicability to these cases where the <decision> part is expected to have a proper decisive force. If a rule is to be attributed deterministic effectiveness. the condition must also assure that the decision will be the best possible at the moment i.e at a given point in discourse. It means that a rule cannot entirely rely on its own judgement and often must consult other rules (refer to Chapter 6 for examples). We do not want to say that the effects of an application of a rule are irreversible. The decisions may get altered at some later point in discourse by the Backward Adjustment Process. Lost links. unobserved forward references. and a simple lack of knowledge are among these factors which may cause revisions to the discourse model maintained by the addressee. Further investigation of above problems leads naturally to another research direction. Among these. formal methods of constructing and evaluating discourse model. tracing active context (focus). and cooperating with individual knowledge bases have to be given attention. A formal definition of the Backward Adjustment Process should follow. including rules specifying when and how the process will be activated. Other related problems such as extracting discourse topic. tracing the structure of focusing. and detecting topic shifts have to be pursued as well.

Yet another question is to work out a knowledge representation language at the level SLA_n that would accommodate various discourse representations and provide an efficient mapping onto the ultimate representation of the universe at the level UA_n.

8.1.3. Toward a Complete Stratified Model

Introduction of the Stratified Model in Chapter 3 generated more questions regarding the Model's final appearance than we could provide answers to in this thesis. In particular. the following problems are left open:

(1) What is the adequate (or alternatively. the minimal) number of strata required?

(2) What can be found at pre-SLA_{n-2} levels on the Left Side of the Model and what transformations are needed there?

(3) How the universe U could be modelled to obtain appropriate approximations UA_i. $i=1 \cdots n$. and what is the nature of the transformations on the Right Side of the Model?

(4) What could be said about the mapping M at the inner-most stratum $<SLA_n, UA_n>$?

Some aspects of these questions have already been addressed elsewhere in this thesis and we do not intend to repeat them here. They all require additional research. We can be more specific about (1) however. According to our Stratified Model hypothesis. the number of strata can be determined by the number of levels at either side of the Model. Except the source language level $SL=SLA_0$ which is taken for granted. we referred explicitly to three representation levels SLA_{n-2} of CAT p-markers. SLA_{n-1} of discourse prototypes in Λ. and SLA_n of discourse representations. This makes $n \geqslant 3$. On one occasion (section 4.10) we assumed existence of the level SLA_{n-3} which appeared to be different than SL. We then get $n \geqslant 4$. Is the four-level Stratified Model a feasible hypothesis? It may appear so if we consider SL as written language. Indeed. it appears reasonable to regard $F_{n-3}=F_1$ as a primarily lexical transformation which produces lexically disambiguated texts at $SLA_{n-3}=SLA_1$. A basic require-

ment here is that the representation at SLA_1 must be directly manipulable by CAT. We believe that future research into questions (2) and (3) will eventually provide a more definite answer to this question. The research into question (4) should verify the other end of our hypothesis that SLA_n is indeed the ultimate discourse representation level in the Model.

8.1.4. Toward a Computer Implementation

The discussion of Chapter 7 seems to indicate that a practical implementation of the Stratified Model is not imminent. The initial effort in this area should concentrate on a computer realization of the F_{n-1}/F_n pair of transformations and development of a knowledge representation language at the level SLA_n. The following work schedule may be considered:

(1) operational implementation of Rules 1 to 12 with static representation of context;

(2) realization of the recursive processing formula for F_{n-1} as defined in section 4.10;

(3) operational implementation of the text cohesion rules of the transformation F_n;

(4) implementation of discourse model active context and primitives to manipulate them;

(5) operational implementation of F_{n-1}/F_n as a pair of cooperating processes.

Realization of subgoals (1) and (2) should not pose much of a problem. An exception may be implementing remote reference rules which have to be given special attention due to their potentially complex impact on the discourse model structure. The remaining points have to be preceded by the theoretical investigation outlined in 7.1.2 above. This limited implementation may be further augmented by appending a categorial grammar parser F_{n-2}. A new cooperation scheme between the three transformations has to be worked out to replace that of (5).

8.2. Conclusions

This dissertation introduces a computationally oriented framework for processing and understanding natural language that is intended to bridge the gap between formal theories of language and Artificial Intelligence practice. The Stratified Model hypothesis derives from the rich theoretical background in philosophy and logics while addressing problems of practical computer realization. It was not our intention in this research to create the complete Stratified Model, nor to testify the Model's psychological credibility. Instead, we consider the Model as a convenient paradigm for investigating more fully the complexity of natural language understanding.

The work reported in this thesis concentrated on selected problems of processing natural language discourse and representing meaning of discourse content. Within the framework provided by the Stratified Model this research has identified domains of three transformations F_{n-2}, F_{n-1}, and F_n, and their respective meaning representation levels SLA_{n-2}, SLA_{n-1}, and SLA_n. The transformation F_{n-2} has been given a largely syntactic character. We assumed that a parsing system based on the categorial grammar CAT defined in Chapter 2 accounted for a part of the transformation F_{n-2}, even if in practice we may need some more sophisticated grammar. Nevertheless, the discourse representation provided by CAT at the level F_{n-2} has subsequently been used to define the transformation F_{n-1}.

The transformation F_{n-1} takes a discourse representation from the level SLA_{n-2} onto the level SLA_{n-1} by computing various inter-sentential dependencies. Thus the SLA_{n-1} representation of discourse consists of a collection of discourse prototypes corresponding to different ways of linking utterances from SLA_{n-2} into partially connected "texts". One of the major contributions of the transformation is a new, uniform approach to the problem of text cohesion by definite descriptions and pronominal references. The transformation operates within a simple recursive formula having at its disposal the notions of context-setting sentence S_1 (a generalized notion of context), "current" utterance S_2, and access to the individual knowledge base KB. We distinguished a number of non-trivial reference situations between S_2 and S_1 involving both referential and non-referential uses of nominal

phrases These included perfect and imperfect contexts, attitude report contexts, conditional contexts, and creative contexts. Our findings have been formalized into translation rules (Rules 1 to 10) which created a fragment of the transformation F_{n-1}. These rules have been incorporated into the transformation's recursive formula. In this environment the rules can be used for processing direct anaphoric references as well as indirect and forward references. The initial scope of the transformation F_{n-1} has been subsequently extended beyond singular terms in the Theory of Names and Descriptions (TND). The concepts of superobject and coordinate allowed for the uniform representation of linguistic terms referring to non-singular objects in the universe, including mass objects and intensional objects. The theory shows that the philosophical notion of intension as a function over possible worlds is not indispensable, and thus offers promise for a computational treatment of intensionality. The TND theory contributes to the problem of inter-sentential dependencies by adding two remote reference rules (Rules 11 and 12) which account for cohesive links between linguistic terms referring to objects classified into different naming levels: supercontext references and subcontext references.

Discourse prototypes generated at the level SLA_{n-1} are subject to processing by the transformation F_n which we presently consider as the final language processing step before mapping onto a universe model. This transformation has to select a single † discourse prototype, the one most accurately corresponding to the speaker's intended message, ‡ restore its full internal connectivity, and produce the final discourse representation at the level SLA_n. Chapter 6 discussed a number of factors which would influence selection of proper cohesive links in discourse, and outlined the framework for developing appropriate translation rules. We approximated the range of applicability of each factor, and classified them according to their origin and decisive impact. This classification, that differs considerably from accounts given in related research, was also intended to facilitate mutual cooperation of the transformation F_n with those preceding it in the Stratified Model. To further promote parallel processing we equipped F_n with the concept of discourse model, active context (focus), and access to

---

† See Chapter 6, however, for more discussion
‡ This attitude is relative to the addressee's ability to understand the message.

individual knowledge base. We defined the process of backward adjustment to allow for retrospective revisions to discourse model and for recovering so-called lost links. i.e. those unobserved at SLA_{n-1} We also discussed the role of discourse topic and topic entity in selecting proper antecedents for definite descriptions and pronominals. The level $<SLA_n, UA_n>$ is yet to be developed. but as we arrive at the ultimate stratum we are far better prepared to build a comprehensive natural language understanding system than Montague was. We have discourse prototypes where Montague had representations of sentences: we have finite functions or finite approximations of continuous functions rather than intensions involving unspecified number of points of reference. At last we do not have to deal with intensional logic what at least does not preclude a possibility of devising a semantic system that would prove to be no more complex computationally than first order logic.

Finally. we summarize computational aspects of our theory in Chapter 7. We argue that the three transformations we discussed are indeed computationally tractable. and that the concept of the Stratified Model for natural language processing promises practical computer realization.

The Theory of Stratified Meaning Representation is at an early stage of development. Many concepts introduced here require further research. Other problems remained virtually untouched at the present time including the problems of discourse structured organization. and they have to be given attention in future. Nonetheless, we believe that this thesis offers an attractive line of research in Artificial Intelligence toward a better understanding of the problems of discourse analysis. and ultimately the phenomenon of natural language understanding in general.

References

Ajdukievicz, K. (1935). "Syntactic Connection." In S. McCall (ed.) *Polish Logic.* Oxford University Press. 207-231. English translation. 1967

Allen, J. F., C. R. Perrault (1980). "Analysing intention in utterances." *Artificial Intelligence.* 15(3).

Barendergt. H. P. (1981). *The Lambda Calculus, its Syntax and Semantics. Studies in Logic. and Fundations of Mathematics.* vol. 103. North-Holland.

Barwise J., J Perry (1983). *Situations and Attitudes.* The MIT Press.

Bennett. M. (1976). "A Variation and Extension of a Montague Fragment of English." In B. H. Partee (ed.). *Montague Grammar.* Academic Press pp. 119-164.

Bobrow. D. G., T. Winograd (1977). "An Overview of KRL. a Knowledge Representation Language." *Cognitive Science* 1. 3-46

Bolc L., T. Strzalkowski (1982). "Transformation of Natural Language into Logical Formulas." *Proceedings Ninth International Conference on Computational Linguistics (COLING).* Prague. North-Holland. 29-35.

Bolc L., T. Strzalkowski (1983). "The Automatic Transformation of Medical Texts into a Deductive Data Base." *Linguistica Computazionale, 3. Proceedings: 7th International Symposium of the Association for Literary and Linguistic Computing (ALLC).* Pisa. 1982. 31-38.

Bolc. L., T. Strzalkowski (1984). "Natural Language Interface to the Question-Answering System for Physicians." *Computers and Artificial Intelligence.* 3(1). 31-46. VEDA. Bratislava.

Bolc. L.. A. Kowalski. M. Kozlowska. T. Strzalkowski (1985). "A Natural Language Information Retrieval System with Extensions towards Fuzzy Reasoning." *International Journal of Man-Machine Studies.* 23. 335-367.

Brachman. R. (1979). "An introduction to KL-ONE." In Research in Natural Language Understanding Annual Report. Bolt Beranek and Newman. Research Report 4785 Cambridge. Mass 13-46.

Brachman. R.. R. E. Fikes. H. J. Levesque (1983). "KRYPTON: a functional approach to knowledge representation." *IEEE Computer.* 16 (10). 67-74.

Brown. G.. G. Yule (1983). *Discourse Analysis.* Cambridge Textbooks in Linguistics Cambridge University Press.

Burton. R. J. S. Brown (1976). "Semantic Grammars: A Technique of Constructing Natural Language Interfaces to Industrial Systems." BBN Report No. 3587. Bolt Beranek and Newman Inc. Cambridge. Mass.

Burge. T. (1975). "Reference and Proper Names." In Davison & Harman (eds.) *The Logic of Grammar.* Dickerson 200-209.

Carlson. G. N. (1978) *Reference to Kinds in English.* Doctoral Dissertation Reproduced by the Indiana University Linguistic Club. Bloomington. Indiana

Carlson. G. N. (1982). "Generic Terms and Generic Sentences." *Journal of Philosophical Logic.* 11 145-181.

Carnap. R. (1947). *Meaning and Necessity.* University of Chicago Press.

Cercone. N.. M. Krause and J. Boates (1983). "Minimal and Almost Minimal Perfect Hash Function

Search " *Computers and Mathematics with Applications,* 9(1) 215-232

Cercone N. J. Boates and M. Krause (1985). "An interactive system for Finding Perfect Hash Functions " *IEEE Software,* 11. 38-53.

Cercone N. R Hadley. F Martin. P. McFetridge. T. Strzalkowski (1984). "Designing and Automating the Quality Assessment of a Knowledge-Based System:The Initial Automated Academic Advisor Experience." *Proceedings: The IEEE Workshop on Principles of Knowledge-Based Systems.* Denver. Colorado. 193-204.

Cercone N. G McCalla (forthcoming). "Accessing Knowledge through Natural Language." To appear in *Advances in Computers.*

Cercone N. T Strzalkowski (1985). "From English to Semantic Networks." In L. Bolc (ed.). *Translating Natural Language into Logical Form* Springer. In press.

Chomsky. N (1957) *Syntactic Structures* Gravenhage. Mouton

Chomsky N (1965) *Aspects of the Theory of Syntax.* Cambridge. Mass.. MIT Press.

Chomsky N (1981) *Lectures on Government and Binding.* Foris. Dordrecht.

Cohen P R (1978). "Planning Speech Acts." TR 118 Department of Computer Science. University of Toronto.

Cohen R (1984) "A Computational Theory of the Function of Clue Words in Argument Understanding." *Proceedings: Tenth International Conference on Computational Linguistics (COLING).* July 1984. Stanford. CA. 251-258.

Colmerauer. A (1978) "Metamorphosis Grammar." In L. Bolc (ed.). *Natural Language Communication with Computers.* Lecture Notes in Computer Science. 63. Springer.

Cooper R (1985) "Meaning Representation in Montague Grammar and Situation Semantics

   Speech given at the Workshop on The Theoretical Approaches to Natural Language

   Understanding. Halifax. Nova Scotia. 28-29 May. 1985

Cresswell. M. J. (1970) "Classical intensional logics" *Theoria*. 36 347 372

Cresswell. M. J (1972) "Intensional Logics and Logical Truth" *Journal of Philosophical Logic* 1

   pp 2-15

Cresswell. M. J (1973) *Logics and Languages* Methuen & Co

Dahl. O (1975) "On generics." In E. L. Keenan (ed) *Formal Semantics of Natural Language*

   Cambridge University Press 99-111

Dahl. V (1985) Personal communication

Delacruz E B (1976) "Factives and Proposition Level Constructions in Montague Grammar" In B

   H Partee (ed) *Montague Grammar* Academic Press

Donnellan K (1971) "Reference and Definite Descriptions" In D D Steinberg L A Jakobovits

   (eds) *Semantics* Cambridge University Press 100 114

Dowty. D R (1976) "Montague Grammar and Lexical Decomposition of Causative Verbs" In B H

   Partee (ed). *Montague Grammar* Academic Press

Dowty. D. R.. R E Wall. S Peters (1981) *Introduction to Montague Semantics* Reidel Dor

   drecht.

Fahlman. S. (1979). *NETL: A system for representing and using real world knowledge* The

   MIT Press

Fillmore. C. (1968) "The case for case" In E. Bach. R. T. Harms (eds.). *Universals in Linguistic Theory* Holt. Rinehart and Winston. 1-88

Frege. G (1892) "On Sense and Reference" In P. T. Geach and M. Black (eds.). *Translations from the Philosophical Writings of Gottlob Frege* Blackwell. Oxford. English translation 1960 56-78

Gazdar. G (1981) "Phrase Structure Grammar" In G. K. Pullum. P. Jacobson (eds.). *On the Nature of Syntactic Representation* Reidel. Dordrecht

Georgeff. M (1984) "Transformations and Reduction Strategies for Typed Lambda Calculus." *ACM Transactions on Programming Languages*, 6(4). 603-631

Grosz. B. J. (1977) "The representation and use of focus in a system for understanding dialogues." *Proceedings Fifth International Joint Conference on Artificial Intelligence (IJCAI)*. Cambridge. Mass.. August 1977. 67-76

Grosz. B. J. (1981) "Focusing and Description in Natural Language Dialogues" In A. Joshi. B. Webber. I. Sag (eds.) *Elements of Discourse Understanding* Cambridge University Press. 84-105

Grosz. B. J. C. L. Sidner (1985) "Discourse Structure and the Proper Treatment of Interruptions" *Proceedings Ninth International Joint Conference on Artificial Intelligence (IJCAI)*. Los Angeles. CA. August 1985. 832-839

Heim. I. R. (1982) *The Semantics of Definite and Indefinite Noun Phrases.* Doctoral Dissertation. University of Massachusetts. Amherst.

Hirst. G (1983) *Semantic Interpretation Against Ambiguity.* Doctoral Dissertation. Technical Report CS-83-25. Department of Computer Science. Brown University.

Hornstein. N. (1984). *Logic as Grammar. An Approach to Meaning in Natural Language.* The MIT Press

Kay. M. (1979). "Functional Grammar" *Proceedings 5th Annual Meeting of the Berkeley Linguistic Society* 142-158

Kripke. S. (1972). "Naming and Necessity" In D. Davison. G. Harman (eds.). *Semantics of Natural Language.* Reidel. Dordrecht. 253-355

Lewis. D. (1976). "General Semantics" In B. H. Partee (ed.). *Montague Grammar.* Academic Press. 1-50

Marcus. M. P. (1980). *Theory of Syntactic Recognition for Natural Language.* The MIT Press Cambridge. Mass

McCarthy. J. P. Hayes (1969). "Some Philosophical Problems from the Standpoint of Artificial Intelligence" *Machine Intelligence.* 4. 463-502

Minsky. M. (1975). "A framework for representing knowledge" In P. Winston (ed.). *The Psychology of Computer Vision.* McGraw Hill. 211-277

Montague. R. (1974a) "Pragmatics" In Thomason (1974)

Montague. R. (1974b) "Pragmatics and Intensional Logic" In Thomason (1974)

Montague. R. (1974c) "On the Nature of Certain Philosophical Entities" In Thomason (1974)

Montague. R. (1974d) "English as a Formal Language." In Thomason (1974)

Montague. R. (1974e). "Universal Grammar." In Thomason (1974)

Montague. R. (1974f). "The Proper Treatment of Quantification in Ordinary English" In Thomason

(1974)

Moore. R.. (1981) "Problems in Logical Form" *Proceedings: 19th Annual Meeting of the ACL.*
Menlo Park CA. 117-124.

Nash-Webber. B.. R Reiter (1977). "Anaphora and Logical Form: On Formal Meaning Representation
for Natural Language." *Proceedings: 5th International Joint Conference on
Artificial Intelligence (IJCAI).* MIT.. Cambridge. Mass. 121-131.

Partee. B. H. (1972). "Opacity.. Coreference. and Pronouns." In D. Davison. G. Harman (eds.).
Semantics of Natural Language. Reidel. Dordrecht. 415-441.

Partee. B. H (1976) "Some Transformational. Extensions of Montague Grammar." In B. H. Partee
(ed.). *Montague Grammar.* Academic Press. 51-76.

Polanyi. L. R. Scha (1984). "A Syntactic Approach to Discourse Semantics." *Proceedings: Tenth
International Conference on Computational Linguistics (COLING).* July 1984.
Stanford CA. 413-419.

Quillian. M (1968) "Semantic Memory." In M. Minsky (ed.). *Semantic Information Processing.*
The MIT Press. 227-270.

Quine. W V (1960). *Word and Object.* The MIT Press. Cambridge. Mass.

Quine. W. V (1973). *The Roots of Reference.* Open Court. La Salle. Illinois.

Radford. A. (1981). *Transformational Syntax.* Cambridge University Press.

Roberts. R.. I. Goldstein (1977). "The FRL Manual." AI Lab Memo 409. MIT. Cambridge.

'Scha. R. J. H. (1983). *Logical Foundations for Question Answering.* M.S. 12.331. Philips

Research Laboratories. Eindhoven

Schank. R. (1972). "Conceptual Dependency: a theory of natural language understanding" *Cognitive Psychology*. 3(4). 552-631.

Schank. R. (ed.) (1975). *Conceptual Information Processing*. North-Holland

Schubert. L. K. (1976). "Extending the expressive power of semantic networks." *Artificial Intelligence* 7 163-198.

Schubert. L. K.. R. G. Goebel. and N. J. Cercone (1979). "The Structure and Organization of a Semantic Net for Comprehension and Inference." In N. V. Findler (ed.). *Associative Networks* Academic Press. New York. 121-175.

Schubert. L. K.. F. J. Pelletier (1982). "From English to Logic Context-Free Computation of Conventional Logical Translation " *American Journal of Computational Linguistics*. 8(1). 27-44.

Sidner. C. (1979). "The role of focusing in interpretation of pronouns " *Procedings 17th Annual Meeting of the ACL*. La Jolla California 77-78

Stalnaker. R. C. (1972). "Pragmatics " In D. Davison and G. Harman (eds.). *Semantics of Natural Language*. Reidel. Dordrecht

Strzalkowski. T. (1983). "ENGRA: Yet Another Parser for English." Technical Report 83-10. Department of Computing Science. Simon Fraser University. Burnaby. B.C.

Strzalkowski. T. (1984). "Toward a Proper Meaning Representation for Natural Languages." Working Paper 1. Natural Language Group. LCCR. Simon Fraser University. Burnaby. B.C.

Strzalkowski. T. (1986a). "Representing Conceptual Dependencies in Discourse." *Proceedings:*

Canadian Artificial Intelligence Conference (CSCSI/SCEIO'86). Montreal Canada, May 21-23, 1986.

Strzalkowski, T. (1986b). "An Approach to Non-Singular Terms in Discourse." Procedings: 11th International Conference on Computational Linguistics (COLING). Bonn, West Germany, August 25-29, 1986.

Strzalkowski, T., N. Cercone (1985). "A Framework for Computing Extra-Sentential References." Proceedings: [The Workshop on] Theoretical Approaches to Natural Language Understanding. Halifax, Nova Scotia. 107-116.

Strzalkowski, T. N. Cercone (1986). "A Framework for Computing Extra-Sentential References." Computational Intelligence, in print.

Tarski, A. (1933). "The Concept of Truth in the Languages of the Deductive Sciences." In A. Tarski, Logic, Semantics, Metamathematics. English translation, Oxford 1956.

Thomason, R. (ed.) (1974) Selected Papers of Richard Montague. Yale University Press.

Thomason, R. H. (1976). "Some Extensions of Montague Grammar." In B. H. Partee (ed.), Montague Grammar. Academic Press. 77-118.

Thomason R H (1980). "A model theory for propositional attitudes." Linguistics and Philosophy. 4. 47-70.

Vendler, Z. (1971). "Singular Terms." In D. D. Steinberg, L. A. Jakobovits (eds.), Semantics. Cambridge University Press. 115-133.

Waltz, D. L. et al. (1976). The PLANES System: Natural Language Access to a Large Data Base. Tech. Report T-34, Coordinated Science Lab., University of Illinois, Urbana.

Warren. D. H. D. L. M. Pereira (1982). "An Efficient. Easily Adaptable System for Interpreting Natural Language Queries." *American Journal of Computational Linguistics* 8(3-4). 110-122.

Warren. D. S.. J. Friedman (1982). "Using Semantics in Noncontext-free Parsing of Montague Grammar." *American Journal of Computational Linguistics.* 8(3-4). 123-138.

Warren. D. S. (1985). "Using Montague Semantics in Natural Language Understanding." *Proceedings: [The Workshop on] Theoretical Approaches to Natural Language Understanding.* Halifax. Nova Scotia. 61-68.

Webber. B. L. (1979). *A Formal Approach to Discourse Anaphora.* Doctoral Dissertation Garland

Wegner. P (1968). *Programming Languages, Information Structures. and Machine Organisation.* McGraw-Hill.

Wilks Y. (1975). "Preference Semantics." In E. L. Keenan (ed.). *Formal Semantics of Natural Language.* Cambridge University Press 329-350

Winograd. T. (1983). *Language as a Cognitive Process* vol.1 *Syntax.* Addison-Wesley

Woods. W.. R. M. Kaplan. B. Nash-Webber (1972) "The Lunar Science Natural Language Information System: Final Report." BBN Report No. 2378 Bolt Beranek and Newman Inc. Cambridge. Mass.

Appendix A

Introduction to λ-calculus

λ-calculus is a very general theory of functions regarded as rules (Barendregt 1981). We restrict our discussion here to type-free λ-calculus in which the objects we are studying are at the same time function and argument. In computational terms we say that no distinction is made between programs (functions) and data (arguments). The formalism has found its way into Computer Science due to its computational character. in contrast with those mathematical theories based on Dirichlet principles (that a function is a graph). Many features of programming languages have been inspired by λ-calculus. Those include Algol and Pascal where procedures can be arguments of procedures. LISP in which procedures may be also outputs of procedures. and GEDANKEN which is explicitly founded on λ-calculus. The following discussion, abbreviated from (Berendregt 1981) and (Wegner 1968). gives an account of fundamental properties of the formalism.

The formal (type-free) λ-calculus deals with functions and their relative behavior. The primitive operation of this theory is therefore **application**. For our purposes we shall say that a function f applies to an argument a and denote it by fa. Whenever we have an expression t which contains a variable x. i.e $t(x)$. we shall call $\lambda x.t(x)$ the function that assigns to an argument, say a, the value $t(a)$. In other words

$$(1) \quad (\lambda x.t(x))a = t(a)$$

If we observe then that any function of several variables can be reduced to unary functions † then we have just learned one of the key principles in λ-calculus known as **λ-abstraction**. From our

---

† $f(x\ y)$ reduces to $f_x = \lambda y.f(x\ y)$ and $f_y = \lambda x.f(x\ y)$

perspective, the most important concepts in λ-calculus are: λ-terms, and the notions of λ-conversion and λ-reducibility.

Definition 1 (Berendregt 1981)

λ-terms are words (expressions) over the following alphabet:

(i) $v_0, v_1 \cdots$ - variables.

(ii) λ - abstractor.

(iii)   (.) - parentheses. □

Definition 2 (Berendregt 1981)

The set of λ-terms is the least set Λ such that

(i) $x \in \Lambda$

(ii)   if $M \in \Lambda$ then $(\lambda x.M) \in \Lambda$

(iii)   if $M, N \in \Lambda$ then $(MN) \in \Lambda$

where x denotes an arbitrary variable. □

In (ii) λx is called the **bound-variable part** of a λ-expression, while M is a body of λ-term. In (iii) M is called an **operator part**, and N an **operand part** of a λ-term. Some examples of λ-terms are given below:

(1) x

(2) (xy)

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

(4) $\lambda xy.(yx) = (\lambda x\,(\lambda y\,(yx)))$

(5) $\lambda xy.(yx)(\lambda z.z) = (\lambda x\,(\lambda y\,((yx)(\lambda z.z))))$

Variables appearing in λ-terms can be divided into three categories: binding, bound, and free.

Definition 3 (Wegner 1968)

(i)   A variable is said to be a **binding variable** if it immediately follows the symbol λ.

(ii)   A given instance of a variable x is said to be **bound** in a λ-term M if it is a binding variable, or if there is a λ-term M' in M of the form $\lambda x.M''$ where M'' includes this instance of x.

(iii) A given instance of a variable x is said to be **free** in λ-term M if it is not bound in M. □

For example in $(\lambda x(xy)\lambda y(xy))$ the variable x is bound in $\lambda x(xy)$ but not in $\lambda y(xy)$.

Definition 4 (Barendregt 1981)

The basic equivalence relation on λ-terms (or λ-convertibility) is defined by following rules:

(1) $(\lambda x.M)N = M_N^x$ wnere M_N^x comes from M by substituting all free occurrences of x by N.

(2) $M = M$.

(3) $M = N \rightarrow N = M$.

(4) $M = N \, \& \, N = L \rightarrow M = L$.

(5) $M = N \rightarrow MZ = NZ$.

(6) $M = N \rightarrow ZM = ZN$.

(7) $M = N \rightarrow \lambda x.M = \lambda x.N$. □

Expressions of the form (MN) are the operator-operand pairs, and specify the application of a function M to an argument N. If M has the form $\lambda x.M$ then the effect of its application to N is the substitution of all free occurrences of x in M by N. The latter rule is known as **λ-reduction** rule and will be denoted by $\rightarrow$ throughout the rest of this section. Below, we give some examples of λ-reductions quoted from (Wegner 1968).

(1) $(\lambda x.x)(yz) \rightarrow (yz)$

(2) $(\lambda x.x)(\lambda x.x) \rightarrow (\lambda x.x)$

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

(4) $(\lambda x.xx)(\lambda y.y)z \rightarrow (\lambda y.y)(\lambda y.y)z \rightarrow (\lambda y.y)z \rightarrow z$ ·

(5) $(\lambda x.(xx))(\lambda x.(xx)) \rightarrow$? yields a nonterminating reduction.

We shall call expressions like (5) **irreducible**. The λ-reduction rule in its unrestricted form, as presented above. may sometimes produce undesired effects. Let us take the λ-term $(\lambda x (\lambda y (xy)))(yz)$. † When λ-reduced. it yields the expression $(\lambda y ((yz)y))$ causing the free occurrence of y in operand part (yz) to be unintentionally unified with the the bound occurrence in operator part. Even more striking example is provided by Wegner (1968) where $(\lambda x (x\lambda x (xy))(uv)) \rightarrow ((uv)\lambda(uv) ((uv)y))$ which is not a well-formed λ-term. It follows that the λ-reduction rule must be constrained somehow to take care of the proper renaming of variables prior to reduction.

Reduction Rule (Wegner 1968)

An expression of the form $(\lambda x.M)A$ can be λ-reduced only if M contains no bound occurrences of x. and A contains no free variables bound in M. □

Renaming Rule (Wegner 1968)

Let M be a fragment of a λ-term other than binding variable. If x is bound in M then M can be replaced by M_y^x provided M contains no free occurrences of x. and y does not occur in M. By M_y^x we mean the term M in which all occurrences of variable x have been replaced by occurrences of variable y. □

· Clearly. two λ-expressions which can be converted into one another by Renaming Rule are equivalent

---

† We shall omit dots in λ-terms whenever it does not lead to ambiguous reading of an expression.

It must be remembered, however, that two equivalent λ-terms are not necessarily reducible to equivalent expressions. It is sometimes said that the value of a λ-term is its **reduced form** i.e. the form from which no further reductions are possible. As we have seen before, not all λ-terms have values. In general, the question whether a λ-term is reducible or not is undecidable. One can easily imagine that for a given λ-term more than one reduction sequence is possible. Thus we are arriving at the important Church-Rosser theorem.

<u>Theorem 1</u> (Church-Rosser)

If a λ-term can be reduced by two different reduction sequences then the obtained expressions are equivalent. □

Unfortunately, the above theorem does not exclude the possibility that a λ-term can have an infinite reduction sequence, even if other sequences yield a reduced form. This situation is illustrated in (Wegner 1968) with the example: $(\lambda x\ (\lambda y.y)((\lambda x\ (xx))(\lambda x\ (xx))))$. If we first reduce by substituting all occurrences of x in $\lambda y.y$ by $((\lambda x\ (xx))(\lambda x\ (xx)))$ then the reduced form of $\lambda y.y$ is obtained. If, on the other hand we first attempt to reduce the operand part of the expression, an infinite reduction sequence will result. It can be shown, however, that the only circumstances in which two different reductions yield finite and infinite reduction sequences for a λ-term is when the λ-term has a form $(\lambda x.M)A$ where A is irreducible and M is reducible and contains no occurrences of x. Proof can be found in (Wegner 1968).

We close our brief presentation of λ-calculus by quoting another important theorem regarding computational properties of the apparatus. The theorem is taken from (Wegner 1968) and its proof can be found therein.

<u>Theorem 2</u>

If a λ-term is reduced from left to right i.e. by applying the leftmost (outermost) reductions first, then the resulting reduction sequence will terminate if and only if the λ-term is reducible. □

This result should be quite clear if one considers what we have said above.