SYNTHESE LIBRARY / VOLUME 253

JAROSLAV PEREGRIN

DOING WORLDS WITH WORDS

Formal Semantics without Formal Metaphysics

SPRINGER-SCIENCE+BUSINESS MEDIA, B.V.

DOING WORLDS WITH WORDS

SYNTHESE LIBRARY

STUDIES IN EPISTEMOLOGY,

LOGIC, METHODOLOGY, AND PHILOSOPHY OF SCIENCE

VOLUME 253

JAROSLAV PEREGRIN
Faculty of Philosophy, Charles University, Prague

DOING WORLDS WITH WORDS

Formal Semantics without Formal Metaphysics

SPRINGER-SCIENCE+BUSINESS MEDIA, B.V.

Library of Congress Cataloging-in-Publication Data

Peregrin, Jaroslav.
Doing worlds with words : formal semantics without formal metaphysics / by Jaroslav Peregrin.
p. cm. -- (Synthese library ; v. 253)
Includes bibliographical references and indexes.
ISBN 978-90-481-4618-5 ISBN 978-94-015-8468-5 (eBook)
DOI 10.1007/978-94-015-8468-5
1. Language and logic. 2. Semantics. 3. Metaphysics. I. Title. II. Series.
BC57.P37 1995
121'.68--dc20 95-36953

ISBN 978-90-481-4618-5

Printed on acid-free paper

for my wife, my lover and my friend (all in one)

The trouble is that the notion of fitting the totality of experience, like the notions of fitting the facts, or of being true to the facts, adds nothing intelligible to the simple concept of being true.

Davidson (1974, pp.193-194)

To construct a theory is, in its most developed or most sophisticated form, to postulate a domain of entities which behave in certain ways set down by the fundamental principles of the theory, and to correlate - perhaps, in a certain sense to identify - complexes of the theoretical entities with certain non-theoretical objects or situations.

Sellars (1956, p.181)

Formalisierungen machen doch ganz offenbar erst Sinn, wenn man mit ihrer Hilfe *inhaltliche Fragen* besser (übersichtlicher) beantworten kann.

Stekeler-Weithofer (1986a, p.388)

CONTENTS

Preface

This book will be read in a number of different ways by different kinds of readers, and I cannot pretend to anticipate all of them; I write from the perspective of a full-time linguist who is a part-time philosopher with a background that (sometimes uneasily) combines a Chomskyan and a Montagovian perspective, with a certain amount of loss of sleep suffered in occasional struggles with the foundational questions that such an attempted synthesis can provoke. But I am quite sure that while different readers relate to the book in different ways, one of its strengths will turn out to be the range and diversity of readers for whom it has something of interest to offer.

I have been acquainted with various temporal stages of this book for about two years now, and am happy to be given an opportunity to say a word or two about it as it is about to go into production. Its author has been surprised that I in fact like his book very much, surprised because on one level the book is an attack on some of the foundations of contemporary model-theoretic semantics, of which I am a committed practitioner. It has been interesting to think about whether and how I can agree with the arguments of this book and still keep doing formal semantics in pretty much the same way (at least with respect to the foundational issues addressed in this book) that I have been doing for a couple of decades; I hope other readers will enjoy the sometimes mind-bending exercise of wrestling with such questions as much as I have.

Peregrin's is a fresh voice coming from a perspective that blends several distinct philosophical traditions and much original thought, on a family of problems that lie at the heart of contemporary semantics, philosophy of language, logic, and foundations of cognitive science. What makes this book unique in my experience is that on the one hand the basic message is in certain ways arguing against important aspects of what are usually taken to be the foundations of contemporary model-theoretic semantics – Montague grammar and its descendants and relatives – but on the other hand the book is written with complete sympathy for not only the aims but most of the practices of model-theoretic semantics. So it should be read, and seems to me could be read (with pleasure as well as profit and without feeling "under attack") by formal semanticists, by philosophers involved in philosophy of language, philosophy of logic, and metaphysics, whatever their current position is with respect to these controversial issues concerning the foundations of semantics. There has tended up to now to be an almost unbridgable gulf: those who attack the foundations of formal semantics tend to be sceptical about the enterprise, and those who are involved in or appreciative of the enterprise are rarely receptive to critics who consider the whole enterprise ill-conceived, who seem not to appreciate it at all, and who place little value on its achievements. Peregrin has bridged that gap in a remarkable way; he might be described as a "friendly critic", in effect trying to clarify what "our" foundations really are, not so that "we" will stop doing what we've been doing, but so that we can get a better understanding of what we're doing, particularly so that we might reexamine what assumptions are essential and what are inessential and perhaps should be modified or replaced. The author examines an interwoven network of ideas that have been inherited by contemporary formal semanticists and philosophers of language from some of our influential forebears, ideas about logic, model theory, and metaphysics, about the relation between natural languages and formalized languages, and about the relations among the world, models of the world, and the interpretation of language, and argues that some of the assumptions in this network are untenable and should be replaced, particularly assumptions concerning what can be explained in terms of what. A most constructive aspect of the book is the way in which replacements are proposed so that the resulting network or fabric is strengthened rather than simply torn apart. The working semanticist can in the end proceed with her work in ways that are largely compatible with or natural improvements upon past practice, if the foundations of the enterprise are understood in the way Peregrin argues that they should be.

There are parts of this book which gave me the experience of "shocks of recognition", as if I was understanding explicitly for the first time things that I could recognize as being right but had never seen that way before. And I think Peregrin has absolutely succeeded in writing in a way that gives full respect to the positions he disagrees with, and full generosity of spirit to those who (like many linguists) have been working as formal semanticists with a possibly unexamined inherited view of what the foundations are. Peregrin addresses the issues in such a way that none of us have to feel "dumb" about accepting or having accepted one or another view of what the nature of the enterprise is. And those who have not particularly worried about the philosophical underpinnings of the semantic enterprise will find this book an excellent and balanced introduction to the problems and the positions that have been taken on them.

One of the appeals of the book for semanticists who may not themselves be philosophers is the constructive way that Peregrin engages in dialogue with people who are on the one hand intrigued by and concerned with the philosophical foundations of the field but even more influenced by the obvious fruitfulness of contemporary formal semantics and therefore not about to simply abandon it no matter what the foundational problems. He has done a beautiful job of showing how one might radically revise the foundations without necessarily having to change much of the superstructure – that's very welcome. In fact for linguists, the book may lead to the interesting realization that the "linguist's outlook" has probably always been more different than we may have realized from that of philosophers like Montague whose research program we have in part adopted and many of whose foundational assumptions we have been at least tacitly agreeing with. A readiness to be quite content with studying "structures" and a tendency to be an "instrumentalist" with respect to metaphysical questions is probably a difference between linguists and philosophers that can be added to the list that starts with which journals we are embarrassed not to have read (Bach 1986, cited by Peregrin in the book).

The book is also a great pleasure to read. Although I am no literary critic, I venture to suggest that over the course of its rewritings it has become a work of art as well as of philosophy. Certainly several of the similes and metaphors are fresh and fun. Perhaps this is one of several ways in which Peregrin has been influenced by Wittgenstein, or perhaps here we are simply seeing some of the talent of the author who (under a pseudonym) had an early novel published in Czech by the famous Canadian publishers of "dissident" Czech literature, 68 Publishers in Toronto.

Barbara H. Partee
Prague, July 1995

Acknowledgements

There are two people especially without whom this book could hardly appear; namely Petr Sgall and Barbara Partee. Were it not for Petr Sgall's encouragement and help, I would hardly have set out to write something like this, and were it not for the advice and encouragement of Barbara Partee, I would hardly have developed it as far as I have. Both of them also provided me with numerous comments and suggestions, without which the book could not have taken the shape it has. I am very grateful for all they have done for me - and done despite the fact that I fear neither of them quite shares my interest for all the questions addressed in the book, let alone the answers that the book offers.

From the viewpoint of direct influence, my gratitude belongs to Pirmin Stekeler-Weithofer, whose book *Grundprobleme der Logik* "awoke me from my dogmatic slumber" and caused me to contemplate my previous activities in the field of formal semantics and thereby see that they were simply the result of being held captive by a certain picture, namely by the picture drawn into our minds by modern formal logic. Before encountering Stekeler-Weithofer's book, I had seen philosophy and, in particular, logic as a way towards the firmest of truths, thereafter I realized that logic is a mere tool that can be used both to elucidate and to blur, and that philosophy cannot be made sense of in any other way than as the Wittgensteinian struggle with the bewitchment of our reason by our language.

The ideas presented in the present book originated during my year's stay at the Department of Philosophy of the University of Konstanz, which was kindly sponsored by the Alexander von Humboldt Foundation. There I had the opportunity not only to discuss the problems that bothered me (and that gave rise to the present book) with Pirmin Stekeler-Weithofer personally, but also to meet other people, such as Friedrich Kambartel, Gottfried Gabriel, Andre Fuhrmann, Arnim von Stechow; and discussions with these people helped me greatly in various ways.

The early articulation of my ideas about logical formalization resulting from this stay appeared as a little book, *Words and Worlds*, published by the Philosophical Faculty of the Charles University in Prague in the series 'Studies in Theoretical and Computational Linguistics'. The present book is a much more elaborate exposure of the same themes. However, my later considerations caused me to change the terminology used in *Words and Worlds* in certain respects, the most crucial being the following: whereas there I considered *necessary truth* as a pre-theoretical concept and *logical truth* as its theoretical counterpart, now I speak about *validity* where I spoke about *logical truth*, and I use the term *logical truth* for the most general part of *necessary truth*.

The present book is based on the further development and more detailed elaboration of the ideas presented in *Words and Worlds*. The work on the book has been made possible by grant no. 794/91 of the Research Support Scheme of the Central European University. Various parts of the manuscript have been essentially improved thanks to critical comments from Petr Hájek, Petr Kolář, Ivana Korbayová, Geert-Jan Kruijff, Josef Mlček and Petr Vopěnka.

1 Introduction: Logic and the Problem of Meaning

The words and expressions we employ to make ourselves understood by other people have meaning; and it is their meaning that capacitates their serving as tools of communication. But what is meaning?

Descartes, Locke, and other pioneers of modern philosophical thinking established a view according to which meanings are reflections or representations of things within human minds. "The use of words," as Locke put it, "is to be sensible marks of ideas, and the ideas they stand for are their proper and immediate signification." That such a representational view of meaning still plays an important role is documented by the fact that in this century it has been defended in various forms by thinkers of otherwise such diverse nature as Edmund Husserl, Vladimir Iljič Lenin, Noam Chomsky, or Jerry Fodor.

However, an alternative approach to meaning is also vibrant in this century, thanks particularly to the influence of Gottlob Frege and his followers. According to this approach, meanings are not internal contents of one's consciousness, but rather objects of an external realm of intangible entities. This new brand of platonism was backed by the development of modern formal logic and mathematics - it was formal set theory that made it possible to supplant meanings from the traditional, mysterious world of ideas into the austere and "scientific" world of sets. Nothing must be presupposed, save the existence of the most ordinary things plus the possibility to group them into sets - this by itself guarantees the existence of the vast world of disparate entities eligible as meanings.

According to this view, formal logic is the tool of disclosing the ultimately true structure of language, and set theory - a branch of formal logic itself - is the universal means of capturing meaning. Such a view has emerged from the writings of Bertrand Russell, Rudolf Carnap, the early Ludwig Wittgenstein and especially Alfred Tarski, who made the ultimate step of rephrasing the problem of meaning as the problem of model theory - the mathematical discipline which studies the relationship between languages and set-theoretical structures. Tarski's approach has been accepted and further elaborated by a number of linguists and philosophers, whose approach is characterized by Richard Montague's denial of any substantial difference between natural language and the languages of formal logic.

Formal logic originally ensued from suspending the relationship between natural language and its logical schematizations, the reason for the suspension being the desire to study the internal properties of formal calculi. However, once formal calculi began to be studied independently of their relationship to natural language, they slowly began to appear not merely as schematizations, but rather as languages on their own count; not as reconstructions of natural language, but rather as its alternatives. However, taken in this way, they immediately appeared substantially incomplete: for whereas with natural language expressions it seems basic that they are linked to objects of the external world, expressions of the formal calculi lacked any such links. Here Tarski entered the scene: his model theory appeared to be precisely what was needed to provide extralinguistic entities to which expressions of formal calculi can be linked. Thus the parallelism between natural language and languages of formal logic appeared to be completed; and scholars like Montague began to deny any real difference between the two kinds of languages. Model theory has been related to natural language itself as an explication of its semantics.

Moreover, however opposed to psychologism Frege's approach to semantics was, the "denotational" view of meaning resulting from it is usually no longer contrasted with representationalism. Many semanticians consider their approaches compatible with both "denotationalism" and "representationalism": they do Tarskian model-theoretical semantics claiming that what they do is to reveal the structure of human consciousness. In fact, Tarskian model theory seems to be the meeting point - it seems to be well suited both to the denotational paradigm and to the modelling of the representative capacities of consciousness. The explicit intention to reconcile these two views of meanings is characteristic of at least two influential approaches to formal semantics of the eighties (*viz* Kamp, 1981; Barwise and Perry, 1983).

Formal semantics, including both the orthodox Fregean one and the one amalgamated with representationalism, often seems to require a formal metaphysics as its foundation. We must, it is suggested, first capture the world - or the way human subjects perceive it - in terms of a set-theoretical construction, and then we simply pair expressions of language with proper elements of the construction. Considered in this way, model-theoretic semantic theory appears to be able to *explain* language; to explicitly point out the entities for which expressions are mere substitutes. We can, it seems, explain the fact why we call snow *white*: well, it is, as Tichý (1988, p.6) stresses, because snow *is* white (or, alternatively, because we perceive it as white)!

One of the main tasks of this book is to show this view to be doomed. I think that the case against metaphysical speculations was made out quite clearly in the pioneer days of formalization of language by Frege, Russell, Carnap and Wittgenstein. In fact, formalization of language, and the logical analysis of language, was originally largely conceived of as a (deadly) weapon against metaphysics. Against deliberate speculations about the structure of the world there was the austere analysis of the structure of language (which proved to determine the structure of the world) leaving no room for speculations. Curiously enough, the weapon has been rapidly tamed and turned into an effective tool of metaphysics: logical analysis of language, which was to replace metaphysics, has provided tools for new, formal metaphysics; and soon it itself has come to be considered a mere spin-off of metaphysics.

However, there is still a third way of perceiving the matter of meaning, a way which is the real *novum* of this century, and this way avoids metaphysics. It is the late-Wittgensteinian view that to be significant or meaningful means to be incorporated into a certain context, especially to play a certain role in a language game. Many model-theoretically oriented semanticians consider this view as giving up explicit semantics, let alone formal semantics. I think that this is not true and one of the aims of this book is to show that it is possible to do formal semantics without falling into metaphysics.

Thus, this book is an attempt to challenge the alliance between (formal) semantics and (formal) metaphysics by pointing out that the Wittgensteinian stance offers a basis for making sense of logic and semantics without succumbing to metaphysics. We put forward the view that logical analysis of language should be seen not as directly modelling the words-things relationships and explicating meaning of a word by pointing out which thing it stands for, but rather as articulating criteria characterizing semantically relevant aspects of language and thereby capturing the unity of meaning in the multiplicity of surface forms.

The metaphysician believes that by pointing out the relevant part of reality, or its counterpart in a model structure, he can give the meaning of an expression, or a ground for the truth of a statement. Given a statement, such as *Aristotle is mortal*, he believes to locate the elements of the world, say the individual Aristotle and the property of being mortal, the junction of which makes up the meaning of the statement and whose occurrence makes the statement true. He is thus willing to consider a pronouncement such as *the individual Aristotle instantiates the property of being mortal* (or, more likely, its logical schematization) as revealing something behind the statement, something that is capable of explaining its truth.

The approach advocated here, on the other hand, results from the conviction that there is no theoretical way of going beyond language, because every theoretical presentation of meaning must inevitably itself be couched in a language. It rejects the meaningfulness of considering pronouncements like *the individual Aristotle instantiates the property of being mortal* as something more than more or less cumbersome paraphrases of the statement the meaning of which they purport to give; accepting, however, that certain paraphrases, if carried out systematically for the whole of language (or for its nontrivial part), may significantly contribute to our understanding language, in that they help us see a relevant structure.

In this sense, this book aims at a criticism of mistaking paraphrases for explications; and especially of the view according to which the problem of meaning can be captured in terms of model theory - as attempted by Tarski, Carnap, Montague, Barwise and other scholars engaged in formal semantics. The criticism aims neither at formal logic, nor at formalization of natural language semantics; it aims at the misuse of formal logic and at the very idea of formal semantics as based on formal metaphysics; and at the mistaking of answers to the questions of empirical linguistics (or psychology or any other empirical science) for answers to the questions of philosophy. Since we see the misuse as the result of long-termed misconceptions about the nature of formal logic, we deem it necessary to trace them back to their roots and to indicate how they could be avoided.

The plan of the book is as follows.

In Chapter I we start by considering the very idea of logical schematization of language. We point out that the basic task of logic can be seen as making a "catalogue" of instances of consequence, and that logical symbols should be considered from the point of view of their contribution to coping with this task. We distinguish between two essentially different modes of employing symbols: the regimentative mode, aimed at suppressing those aspects of language which are not considered relevant, and the abstractive mode, aimed at articulating types of expressions and statements. In the last section of the chapter we articulate the difference which we consider crucial:

the difference between logical symbolization, the taking of symbols merely as handy shortcuts for language expressions (remaining connected with the expressions by an "umbilical cord"), and logical formalizations, the granting to symbols of an independent life of their own.

In Chapter II we consider ways to catalogue types of schematic representations of instances of consequence, or of necessary truths. We point out that although it is not possible to simply list them, it is possible to conceive of their "potential list", via rules that would generate such a list - an idea which leads directly to the concept of axiomatic system. This leads us to conclude that the basic sense of an axiomatic system is its ability to provide a finite characterization of an infinite number of necessary truths, that it provides what we shall call a *criterial reconstruction* of necessary truth.

Chapter III begins with the exposition of Tarski's challenge to the adequacy of the proof-theoretic approach to consequence or necessary truth. We try to show that Tarski's proposals entail an essential change in the nature of logic, a change which is hardly compatible with the notion of logic as pursued until then. We argue that the Tarskian model-theoretic approach cannot be considered as a way of reducing the concepts of consequence and necessary truth to the concepts of reference and satisfaction; that the only way to make sense of it is to see it as stating necessary truth in different terms, as attempting, again, at its *criterial reconstruction*. We show that Tarskian model theory can be seen as driven by the idea of a criterion better than the one offered by proof theory; that, however, it cannot be considered as quite successful. In particular, we deny that it makes much sense to consider Tarskian model theory as a "representational semantics" in the sense of Etchemendy (1990).

In Chapter IV we turn our attention to where the nature of the model-theoretical approach comes to the open, to accounting for quantificational statements. We show that model theory is based on the assumption that we can in general reduce the truth of every quantificational statement to the truth of its instances; and we point out that this cannot be done in a nontrivial way. We show that the standard model-theoretical approach has come to rest on the doctrine of "logical atomism", the conviction that the truth value of every statement should be reducible to those of the elementary statements constituting a basis. We indicate that reference can be seen as a means of grasping the basic patterns of consequence present within language.

In Chapter V we consider the general concept of truth. We argue that a logician or a philosopher can offer no account for truth beyond the account of the dependence of truth of some statements on truth of other statements, i.e. beyond the account of consequence. We conclude that the correspondence theory of truth, as put forward by the early Wittgenstein and by Tarski, is closely connected with the "representational" view of logic and that it is hence doomed. We reject the idea that the concept of truth should and could be reduced to the concept of correspondence and we conclude that the only theoretical account which can be given for truth can again be seen as its criterial reconstruction.

In Chapter VI we propose that to get from the concept of reference to a concept that could provide for a formal reconstruction of the notion of meaning, we must change the concept in two ways. First, we must abandon our seeing the opposition between extralogical and logical expressions as that between those that do refer, and those that do not - we must consider all expressions as acquiring semantical values (some of them fixed, other variable). Second, we must take into account not only the range of necessary truths accounted for by standard logic, but also necessary truths concerning modal aspects of language - this makes it possible to reduce the theory of necessary truth to a part of the theory of truth *simpliciter* (namely to the theory of truth *simpliciter* of certain modal statements). In this way we turn the concept of reference into the concept of *denotation*.

In Chapter VII we address the concept of meaning. We distinguish between two essentially different concepts of meaning: one notion taking the meaning of an expression as a pre-existing object which gets baptized by the expression, and the other notion taking the meaning as resulting from hypostatizing the way the expression functions in language. We show that the first of these views is closely allied to representationalism, that it leads to the picture of language as a nomenclature (thus amounting to what Quine calls the *museum myth*), and that it fails to provide an adequate account. The second notion, on the other hand, leads rather to the view of language as a structure, as the structure articulated especially by the patterns of consequence. We indicate that in the same way as reference can be seen as a means of reconstructing the most basic patterns of consequence present within language, meaning can be seen as reconstructing the patterns of consequence in their entirety; and that it is the concept of denotation, as developed in the preceding chapter, that is plausible to formally reconstruct the concept of meaning. Thus, we claim that meanings are best looked upon as means of such a reconstruction, as means of capturing unity of sense within multiplicity of surface forms.

In Chapter VIII we analyze the very concept of criterial reconstruction, which we consider plays such an important role for our understanding of the possibilities and limitations of a theory of language. We indicate that this concept is inseparably connected with the concept of infinity, and we analyze the latter concept. We argue that it is plausible to see the infinity present within our world as merely potential (i.e. as brought about by boundless continuability of various processes), and we distinguish between infinite sets that are criterial (their boundlessness is a matter of boundless recursive applicability of explicitly articulated rules) and those that are uncriterial (their boundlessness is a matter simply of our boundless capability of going on with certain activities). We propose to see theory as the struggle for reconstructing the uncriterial, pre-theoretical in the criterial way, thus making the uncomprehensible comprehensible. In particular we propose that to make a theory of language means to criterially (i.e. finitely) reconstruct infinite classes; syntax being a matter of criterial reconstruction of the concept of well-formedness (or of the - potentially infinite - class of well-formed expressions), semantics that of criterial reconstruction of the concept of truth (i.e. of the - again potentially infinite - class of true statements).

In Chapter IX we return to one of the assumptions traditionally underlying the very idea of logic - the idea of the existence of consequence as something absolute and definite - and we consider the challenge posed to it by the well known argument of Quine. We show that many of Quine's results were anticipated by Wittgenstein (and that in many aspects Wittgenstein's results are deeper) and we conclude that although it is no longer possible to consider necessary truth as something absolute, this result does not undermine logic - what it undermines is the metaphysical doctrine of logical atomism which is often unwarrantedly considered as inseparable from logic.

In Chapter X we give a general overview of the nature of language and the language-world relationship; we show how the view of these matters evolved in the present century and how it has instigated the present confusions.

2 Roots of Logical Schematization

2.1 Frege's Program

To answer the question *what is logic* is anything but simple; and it is surprising how few of those who write on logic try to give an explicit and precise answer to it[1].

The majority of logicians and philosophers considering the nature of logic seem to be certain about one thing: that logic should help us sort arguments and proofs into right and wrong, or sound and unsound. This intention is clear already in Antiquity; it is, however, even more manifest when we consider the origins of modern symbolic logic. Frege, the father of symbolic logic, makes the intentions of his *Begriffsschrift* explicit[2]:

> [Die vorliegende Begriffsschrift] soll also zunächst dazu dienen, die Bündigkeit einer Schlußkette auf die sicherste Weise zu prüfen und jede Voraussetzung, die sich unbemerkt einschleichen will, anzuzeigen, damit letztere auf ihren Ursprung untersucht werden könne.[3]

A mathematical theorem is a statement to the effect that something, some *conclusion*, is true whenever something other, some *assumptions*, are true. Let us call the fact that there are statements which cannot be false when some other statements are true *consequence*; any particular instance of consequence will be

---

[1]*Bonmots* of the kind of Lemmon's (1965) "the best way to find out what logic is, is to do some" may be illuminating, they can, however, hardly substitute a real analysis of the nature of logic.

[2]Frege (1879, p.IV).

[3]Its [of the present ideography] first purpose, therefore, is to provide us with the most reliable test of the validity of a chain of inferences and to point out every presupposition that tries to sneak in unnoticed, so that its origin can be investigated.

called a *(valid) inference*. Thus it is valid to infer (3) from (1) and (2) (i.e.: it holds that (3) cannot be false whenever (1) and (2) are true); and it is also valid to infer (4) from (1) and (3).

Aristotle is human (1)
All humans are mortal (2)
Aristotle is mortal (3)
Aristotle is human and Aristotle is mortal (4)

A mathematical theorem thus states an inference (sometimes with an empty set of assumptions, and sometimes with assumptions not explicitly articulated, but only tacitly presupposed). The proof of such a theorem consists in the decomposition of the inference into a chain of "simple" inferences (*Schlußkette*). What Frege urges is that it must be made transparent that each of the inferences of the chain links up with the preceding ones. (It goes without saying that the inferences the proof consists of must be simple enough to be immediately recognizable as valid.) If, for example, we take both the inferences mentioned above, i.e. the inference of (3) from (1) and (2) and that of (4) from (1) and (3), for simple (which appears to be quite plausible), then we can prove the "theorem" stating the inference of (4) from (1) and (2) by pointing out that it can be decomposed into the two-element chain consisting of the two simple inferences.

What is in question here is thus the problem of *classification*, of the distinguishing between valid inferences and invalid ones. That there are valid inferences and valid proofs is a fact, *a fact which is constitutive to the very idea of logic* (and we thus can hardly expect that it will be explained by logic, at least if under explanation we understand more than summarization and systematization). If we denied that arguments and proofs are valid or invalid "by themselves" (i.e. prior to any logical theory), then we would deny that there could be a logic. In this sense logic is unequivocally *descriptive*. Proofs can be understood as explaining inferences by their reduction to some other, simpler, or more basic inferences; the inferences which we take as basic must, however, simply be taken as self-evident. That (3) cannot be false as soon as (1) and (2) are true is a plain fact, a fact which cannot be explained, but which at most can be registered.

An inference concerns a group of statements (the assumptions; the *antecedent* of the inference) and a statement (the conclusion; the *consequent* of the inference). The inference is valid iff the statement in its consequent cannot be false whenever all the statements in its antecedent are true. To check the continuity of a chain of inferences means to check if the statements in the antecedent of any involved inference also occur either in the consequent of some of the previous inferences or in the antecedent of the inference being proved.

Since checking continuity is a principally straightforward, routine matter, recognizing validity of any proof may appear to be immediate. It may thus seem that any nontrivial logical theory is superfluous. Things are, however, at this point not so simple: the problem is that the boundary between evidently valid inferences and inferences which can raise suspicion is notoriously fuzzy. The situation is similar to that in the field of law when we want to sort deeds into the acceptable and the reprehensible: we all know in principle which deeds are acceptable and which are not (if we did not know this, then law would be completely wilful and hence meaningless), but we nevertheless need an explicit list of the forbidden deeds to facilitate relatively quick and unambiguous assessment of individual cases. So what logic aims at - primarily - is an equivalent of the code: a kind of catalogue of basic valid inferences. What we need, in other words, is an explicit *criterion of validity of inferences*.

To summarize: *There are valid arguments and proofs (prior to any logical theory!) and there are invalid ones. The business of logic is to provide techniques for distinguishing those which are valid from those which are not, to make valid arguments obviously valid. In other words, logic is to deliver a criterion of validity.* That is, logic is to present, as Hacking (1979, p.290) puts it, "rules for certain transitions between sentences"; the rules being "descriptions, or, perhaps, codifications of what one knows when one knows how to make certain transitions that we call logical".

2.2 Logico-Grammatical Analysis

The parallel between logic and law can be extended: the basic problem is to represent an infinite number of items[4] (the infinite number of all concrete forbidden deeds in the case of law, and the infinite number of all valid inferences in the case of logic) by means of a finite list. How the problem is really handled is clear: in the case of a code we do not list all concrete deeds (that it is forbidden for Aristotle to kill Plato, that it is forbidden for Plato to kill Socrates,...), but rather only *types* of such deeds (that it is forbidden for a human being to kill another human being); in the case of logic we do not list all valid inferences (that *Aristotle is mortal* follows from *All humans are mortal* and *Aristotle is human*, that *Plato is wise* follows from *Plato is a philosopher* and *All philosophers are wise,...*) but rather only *types* of such inferences (that *A is B* follows from *All Cs are B* and *A is C*). However, while it is enough to make use of general formulations to articulate types of deeds, in order to express types of valid inferences in a clear and comprehensible way we need to make use of symbols. We may try to say that the inference of (3) from (1) and (2) is an instance of the general law that "if someone is something and all those somethings are something other, then the someone is the something other", but it is evident that only some kind of symbolization allows us to articulate the inference type clearly.

The origin of logical formalization is the act of replacement of a statement or its part by a symbol. We may replace *Aristotle* and *mortal* in *Aristotle is mortal* by *A* and *B*, respectively, and we gain the schema *A is B* which may then be considered to represent not only the statement *Aristotle is mortal*, but also the statement *Plato is wise* and all the like. What makes logical symbolization into a nontrivial matter is the fact that the parts of statements we replace by symbols need not be the *apparent* parts; to reach an efficient symbolization we need to ascribe to language a part-whole structure more or less different from the apparent one.

The point is that if we take language and its part-whole structure at face value, then we would not be able to arrive at a reasonable articulation of types

[4]We shall see later that a proper understanding of the concept of *infinity* is crucial for understanding the general possibilities and limitations of the formalization of language; at this juncture, however, we use the concept in its common sense.

of inferences. We would, for example, not be able to subsume the inference of *Mickey is mortal* from *Mickey is a mouse* and *All mice are mortal* under the schema stating that *A is B* follows from *All Cs are B* and *A is C*: *Mickey is a mouse* and *All mice are mortal* have no common part for which *C* could be considered to stand. What we need in this particular example is to consider *mouse* and *mice* as two forms of the same item; what we need in general is a much more abstract understanding of the part-whole structure of our language.

The character of the part-whole structure adequate for our language would, of course, depend on the goals for which it is carried out, and in particular on what kind of inferences we are about to summarize. If we restricted ourselves to inferences of the kind such as that of (4) from (1) and (3), then it would be enough to acknowledge that (1) and (3) are parts of (4) (and thus it would be enough to have symbols for whole statements). If we wanted to account for the inference of (3) from (1) and (2), we would have to recognize *human* (or *to be human*) as the common part of (1) and (2) and *Aristotle* as the common part of (1) and (3) (thus requiring special symbols for subjects and predicates); and if, moreover, we wanted to account for (1) being inferable from *Aristotle is a great human*, then we would have to go even deeper: we would have to acknowledge that the predicate of (1) is a part of *to be a great human*.

This means that *an efficient formalization of language is connected with positing some suitable part-whole structure, i.e. with a grammatical analysis*. If this analysis is guided by the intention to summarize inferences, then it can lead to what we can call *logical form*[5]; if our intention is simply the general distinction between the relevant aspects of the structure of the sentence and the irrelevant ones, then we may arrive at some notion of Chomskyan *deep structure*, or at what Sgall duly calls *tectogrammatical representation*[6]. The

[5]As Quine (1980, p.21) puts it: "What we call logical form is what grammatical form becomes when grammar is revised so as to make for efficient general methods of exploring the interdependence of sentences in respect of their truth values." This is what Frege has in mind when continuing the statement quoted above by: "Deshalb ist auf den Ausdruck alles dessen verzichtet worden, was für die *Schlußfolge* ohne Bedeutung ist." ["That is why I decided to forego expressing anything that is without significance for the *inferential sequence*."]

[6]Sgall et al (1986). The term *tectogrammatics* is, however, due to Curry (1961).

first case is what interests us primarily; in this case we shall speak about a *logico*-grammatical analysis[7].

A grammatical (or logico-grammatical) analysis allows us to see expressions in a new way: it allows us to consider them from the viewpoint of their capability of being part of other expressions. This also gives rise to the concept of a *(logico-)grammatical category*: two expressions are said to be of the same grammatical category iff the ways they can be incorporated into other expressions coincide, i.e. if we can freely replace one of them by the other without violating grammaticality.

2.3 Regimentation

As we have just seen, a logico-grammatical analysis can lead to the conclusion that an expression consists of some parts other than the apparent ones. We may, for example, conclude that there is a common part of the statements *Mickey is a mouse* and *All mice are mortal*, and that *mice* and *a mouse* are two different "forms" or "manifestations" of this part. Now we may take a symbol, say **M**, as representing this abstract part and we can write *Mickey is* **M** instead of *Mickey is a mouse* and *All* **M** *are mortal* instead of *All mice are mortal* to make our conclusion explicit.

If we proceed this way to reach such a representation of statements in which all parts are replaced by symbols, then what we are engaged in can be called, in accordance with Quine (1960a), *regimentation*. The aim of regimentation is to make the accepted logico-grammatical analysis explicit: to "reify" the abstract units posited by the analysis and to turn statements, sequences of words, into sequences of what has been acknowledged as their parts. If the accepted logico-grammatical analysis allows for the same parts to be composed in alternative ways yielding different wholes (hence if it allows for alternative "grammatical rules"), then the particular way present in the expression in question should be made explicit by regimentation. We may, for example, decide to consider *Aristotle admires Plato* and *Aristotle is admired by*

[7]Later we shall argue that the difference between the general grammatical analysis and the logico-grammatical one is not an essential one; that if grammatical analysis is carried out with the perspective of capturing meaning, it is bound to be *logico*-grammatical.

Plato as the results of two different ways of combining *Aristotle*, *to admire* and *Plato*, and we may decide to indicate this by the order in which the symbols for *Aristotle* and *Plato* occur in the regimentation.

Regimentation thus results in a "symbolic language" which is to be understood as a reconstruction of (a part of) natural language. The primitive symbols used as means of regimentation will be called *constants*; the complex symbolic expressions built of constants in a manner that makes them correspond to the expressions of natural language will be called *constant symbolic expressions*. Those constant symbolic expressions which regiment statements of natural language will be called *constant symbolic statements*[8].

2.4 Classical Ways of Regimentation

Let us summarize the most traditional ways of regimentation, which found their expression in classical logic. First, some statements are acknowledged to consist of other statements plus certain operators. Thus (4) is considered to consist of (1), (3), and the conjunction operator $\wedge$ (in fact the regimentation of *and*); (5) is considered to consist of (1), (3) and the disjunction operator $\vee$; (6) of (1), (3) and the implication operator $\rightarrow$; and (7) of (1) and the negational operator $\neg$.

Aristotle is human or Aristotle is mortal (5)
If Aristotle is human, then Aristotle is mortal (6)
Aristotle is not human (7)

Statements which do not contain other statements are considered to consist of a verb phrase and one or more noun phrases. Thus (1) is considered to consist of the verb phrase *to be human* and the noun phrase *Aristotle*; (8) is

[8]Under *statement* we understand an expression which can be asserted. We do not consider the concept of statement to be reducible to some simpler concepts, e.g. to grammatical ones (every statement is certainly a sentence, but not necessarily every sentence is a statement); we consider the knowledge of language to consist not only in the ability to tell an expression from a non-expression, but also in the ability to tell an assertible expression (i.e. a statement) from an expression which cannot be meaningfully asserted.

considered to consist of the verb phrase *to admire* and of the noun phrases *Aristotle* and *Plato*.

Plato admires Aristotle (8)

Verb phrases are regimented by *predicates*, names by *terms*; a statement consisting of a verb phrase and names is regimented by the predicate followed by the list of terms in parentheses. If we let the constants **Ar**, **Pl**, **Hu**, **Mo** and **Ad** stand for *Aristotle*, *Plato*, *to be human*, *to be mortal* and *to admire*, respectively, then (1), (3) and (8) will be regimented by (1'), (3') and (8'); then (4) through (7) by (4') through (7')[9].

$$\mathbf{Hu(Ar)} \tag{1'}$$
$$\mathbf{Mo(Ar)} \tag{3'}$$
$$\mathbf{Ad(Pl,Ar)} \tag{8'}$$
$$\mathbf{Hu(Ar)} \wedge \mathbf{Mo(Ar)} \tag{4'}$$
$$\mathbf{Hu(Ar)} \vee \mathbf{Mo(Ar)} \tag{5'}$$
$$\mathbf{Hu(Ar)} \rightarrow \mathbf{Mo(Ar)} \tag{6'}$$
$$\neg \mathbf{Hu(Ar)} \tag{7'}$$

We can now write (9) to indicate that it is valid to infer (4') from (1') and (3').

$$\mathbf{Hu(Ar)},\ \mathbf{Mo(Ar)} \Rightarrow \mathbf{Hu(Ar)} \wedge \mathbf{Mo(Ar)} \tag{9}$$

Terms can be considered to stand not only for proper names, but for other name-like expressions as well. To capture certain kinds of inferences, especially those employed in mathematics, we may further need to acknowledge that a term (e.g. *one plus two*) can consist of an operator (*plus*) and other terms (*one* and *two*).

Moreover, it may seem plausible to extend the range of expressions which terms can be considered to schematize to the whole class of noun phrases, i.e. to include not only *Aristotle* or *one plus two*, but also *everybody*,

[9]Let us stress that the fact that we have chosen symbols which, so to say, wear what they stand for on their sleeves is in no sense essential. Evidently we could have chosen equally well $\mathbf{T}_1$ to stand for *Aristotle*, $\mathbf{T}_2$ for *Plato*, $\mathbf{P}_1$ for *to be mortal*, etc.

nobody, *the king of France* etc. It was Russell who pointed out that this is only seemingly plausible: that as far as logical schematization is concerned, we do better when we have terms only for noun phrases of the former kind, and when we employ quantifiers to account for those of the latter.

Let us consider the classical case of *everybody* and *somebody* to see how the attempt to assimilate them to terms can lead us to the introduction of quantifiers. If we write $\forall$ for *everybody* and $\exists$ for *somebody*, and if consider these new terms as quite on a par with **Ar**, **Pl** and the like, we have to write (10') through (12') as regimentations of (10) through (12).

Everybody is human (10)
Somebody is mortal (11)
Everybody admires someone (12)

$\mathbf{Hu}(\forall)$ (10')
$\mathbf{Mo}(\exists)$ (11')
$\mathbf{Ad}(\forall,\exists)$ (12')

If we restrict ourselves to unary predicates, then this treatment is quite sufficient. It allows us to capture the inference of (1') from (10') and that of (11') from (3').

However, if predicates of greater arities are taken into account, then problems may arise. The statement (12), for example, is primarily understood as claiming that everybody has his own idol, but it seems to be also possible to understand it as claiming that there is one person who is admired by everybody. In this sense, the statement may be considered to be ambiguous between two readings[10], and hence it may be perceived as consisting of its parts in two different ways. The two ways may be indicated by various auxiliary means such as indices or braces[11], but the usual way is to use auxiliary letters. Instead of (10') and (11') we thus have (10'a) and (11'a), and instead of (12') we have (12'a) and (12'b).

---

[10]The ambiguity is quite obvious with respect to statements as *Every philosopher admires one logician*. In spoken language, this ambiguity may be resolved by means of intonation or other "topic-focus articulating" means; see Sgall et al. (1986).

[11]See e.g. the employment of braces by Bourbaki (1958).

$\forall x\mathbf{Hu}(x)$ (10'a)
$\exists x\mathbf{Mo}(x)$ (11'a)
$\forall x\exists y\mathbf{Ad}(x,y)$ (12'a)
$\exists y\forall x\mathbf{Ad}(x,y)$ (12'b)

Using this notation, we can regiment (2) as (2') and we can express the fact that (3') follows from (1') and (2') as (13):

$\forall x(\mathbf{Hu}(x)\rightarrow\mathbf{Mo}(x))$ (2')
$\mathbf{Hu}(\mathbf{Ar}),\ \forall x(\mathbf{Hu}(x)\rightarrow\mathbf{Mo}(x)) \Rightarrow \mathbf{Mo}(\mathbf{Ar})$ (13)

The fact that the regimentation of the noun phrase *somebody* is to be of a character quite different from that of *Aristotle* brings about some complications. While in natural language (more precisely in natural language with a fixed word order) there is only one way to grammatically concatenate *everybody*, *to admire* and *somebody* (in this order), namely (12), we need two such ways on the level of regimentation. If we consider a ternary predicate instead of **Ad**, we need six different ways, and so on. The grammar of the language arising out of regimentation thus becomes a menacing complexity. However, this implausibility can be overcome if we break the direct correspondence between the grammar of natural language and that of the symbolic one: if we start to consider symbols of the kind as x and y of (12'a) and (12'b) as grammatically on par with constant terms. Such a step leads to a drastic simplification resulting in the usual simple grammatical structure of the predicate calculus. But we must not forget to keep apart two essentially distinct things: the grammatical structure of the predicate calculus, which is a purely technical matter, and the grammatical structure of natural language, which is what the predicate calculus is to reflect contentually.

In this way the whole language of the traditional predicate calculus can be considered as a means of regimentation of natural language. Also, alternatives and amplifications of the classical predicate calculus can be understood as a means of regimentation[12].

---

[12]Thus too, the higher-order predicate calculi can be considered to result from the acknowledgment of the fact that a predicate can yield a statement not only together with a term, but also together with another predicate (witness *to be glad is wonderful* as consisting of *to be glad* and *to be wonderful*); and modal calculi to result from the acknowledgment of

Regimentation can be conceived of as the sifting of natural language through the sieve of relevance. What is relevant for the purposes of capturing the instances of consequence for which one wants to account is unambiguously retained in the resulting symbolic representation, and that which is not vanishes. The problem is to provide a sieve fine just enough - if it is too fine, we lose something substantial for our purpose; whereas if it is not fine enough, we do not get rid of all irrelevant idiosyncrasies. However, provided the sieve is as it should be, we need not make a difference between a natural expression and its symbolic regimentation, because the aspects of the natural expression in which it differs from its regimentation are just those which are irrelevant for us. Keeping this in mind we thus sometimes do not distinguish between an expression of natural language and its symbolization and we simply speak about *an expression* or *a statement*.

2.5 Abstraction

Regimentation helps us express inferences in a way suitable for summarization: it is, however, not the summarization itself. Regimentation does not leave the level of concrete inferences; it does not lead to any representation of *types* of inferences. To get further we must introduce symbols functioning in another way than those introduced so far. We need symbols which would not stand for definite (although sometimes abstract) parts of natural language, but rather symbols which stand for "fixed but arbitrary" expressions.

What makes it possible to represent an infinite number of concrete inferences by a single schema is that some expressions can be freely replaced by other expressions of corresponding grammatical categories in some valid inferences without the inferences' validity being disturbed. Thus we can replace **Ar** by **Pl**, or by whatever other terms in (13) and it will still be a valid inference. This means that we can introduce a new symbol, say $\boldsymbol{T}$, and summarize all the inferences so obtainable from (13) by (14):

$$\mathbf{Hu}(\boldsymbol{T}),\ \forall x(\mathbf{Hu}(x)\rightarrow\mathbf{Mo}(x)) \Rightarrow \mathbf{Mo}(\boldsymbol{T}) \tag{14}$$

---

new ways of statements being parts of other statements (*bachelors are necessarily unmarried* as consisting of *necessary* and *bachelors are unmarried*).

To say that (14) is valid is to say that by whatever term we replace $\boldsymbol{T}$ in it, we gain a valid inference.

In the same manner we can further replace **Hu** and **Mo** by other unary predicates; if we use new symbols $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ in the same manner as $\boldsymbol{T}$, we can summarize a wider class of valid inferences by (15). (However, we clearly cannot go on: if we replace $\forall$ by $\exists$, or $\rightarrow$ by $\wedge$, what we gain are no longer necessarily valid inferences.)

$$\boldsymbol{P}_1(\boldsymbol{T}),\ \forall x(\boldsymbol{P}_1(x)\rightarrow\boldsymbol{P}_2(x)) \Rightarrow \boldsymbol{P}_2(\boldsymbol{T}) \tag{15}$$

(14) and (15) result from symbolization of a kind essentially different from regimentation; we shall call it *abstraction*. Probably the least confusing way to label the symbols of the new kind is to call them *parameters*. To articulate (15) we required parametric terms and parametric predicates; and to articulate many elementary types of valid inferences we require also parametric statements; but there seems to be no use for parametric connectives or quantifiers.[13] Thus, abstraction is a matter of only some categories of expressions, not of others.

Unlike constants, parameters do not stand for anything definite in natural language, but are mere "place-holders" used to signify that there is something missing in the place they occupy. "Expressions" containing them, such as those in (14) and (15), are mere matrices which must be "saturated" to yield real (i.e.: constant) expressions. We may call such expressions *parametric (symbolic) expressions*; instead of speaking about parametric statements we shall sometimes, as we have done so far, speak also about *schemata*. Only when

---

[13]Parameters are normally means of logical theory, they are means of talking about unspecified constants. However, here, whilst concurrently employing them in this role, we intend also to *talk about* them. This brings about a certain ambiguity: if we say something like 'If *P* is a unary predicate and *T* is a term, then *P(T)* is a statement', then we simply *use* '*P*' and '*T*', we use them as means of talking about an unspecified constant predicate and an unspecified constant term. Whereas when we say, e.g., 'The parametric statement *P(T)* consists of the parametric predicate *P* and the parametric term *T*', we *mention* '*P*' and '*T*', we talk about them. Where we shall want to stress that we are mentioning parameters, that we are talking *about* them, we shall use bold italics, like 'the parametric statement ***P(T)*** consists of the parametric predicate ***P*** and the parametric term ***T***'. So italic print marks the constant/parameter distinction, and bold print marks the - sometimes vague and sometimes unimportant - mention/use distinction.

parameters are mapped on constants is any parametric expression mapped on some constant expression.

While natural statements, and derivatively constant symbolic statements, can be true or false, schemata are neither true nor false. A schema can be valid or invalid: it is valid if all its instances are true, and it is invalid otherwise. The situation is schematized in the following table.

NATURAL STATEMENT		CONSTANT SYMBOLIC STATEMENT		SCHEMA		SCHEMA
(TRUE/FALSE)				(VALID/INVALID)		
Aristotle is human	———	**Hu(Ar)**				
				P(**Ar**)		
Aristotle is mortal	———	**Mo(Ar)**				*P*(*T*)
				Mo(*T*)		
Plato is mortal	———	**Mo(Pl)**				
	REGIMENTATION		ABSTRACTION		ABSTRACTION	

Table 1

By introducing constant symbolic expressions we have gained a symbolic language. After now having introduced parameters, we may consider *another* language: a language differing from the previous one in that all its terms, functors, and predicates are not constants, but rather parameters. The new language differs from the original one in that its statements cannot be said to be true or false, but rather valid or invalid; the structure of both languages is the same, the only difference is how they are understood. This means that we have two alternative ways to understand the predicate calculus and, in general, any other symbolic language.

2.6 Further Flights of Regimentation

Regimentation, as exposed so far, simply means letting all those aspects of expressions that are irrelevant from the point of view of capturing consequence go by the board. Only the relevant skeleton remains. However, there are various reasons which may lead us to further rearrange the bones of the skeleton.

Via the process of regimentation we arrived at certain categories of expressions; the way of standard logic we adopted led us to the categories of statements, terms, predicates, one- and two-place operators and quantifiers. In contrast to the abundant supply of grammatical categories of natural languages, this set of categories seems to be about the minimum needed to express our talk about the world - hence it may be considered as something amounting to *ontology*. Some of the categories, namely the categories of operators and quantifiers contain only a small number of expressions, others, namely those of terms and predicates, must be considered to contain large stocks of primitive items. Thus, it is the latter that may be considered to reflect the really ontologically basic kinds of entities, namely those of objects, properties and relations. The basic syntactic opposition between term and predicate is hence considered to amount to the basic ontological opposition between object and concept[14].

Now we may consider the question whether this stock of basic expressions - or the corresponding stock of ontological primitives - is really the minimal one, or, as Russell (1914, p.51) puts it, whether it is really "the smallest store of materials with which a given logical or semantic edifice can be constructed". We may, for example, want to see whether it would be possible to decrease the stock of properties and relations - for whereas the reality of objects seems to be straightforward, the reality of properties and relations has a far more problematic nature (as documented by the eternal realist-nominalist controversy). Translated into the terms of logical schematization, we may want to dispense with as many predicates as possible (even at the costs of introducing new terms).

One way to do this is, of course, to make predicates into terms and to keep only predicates amounting to the relation of instantiation of a predicate-object by an object proper. Let us introduce the term T_P for every predicate P and let us restrict ourselves to the predicate $\mathbf{INST}_n$ for every natural n; and let us always write $\mathbf{INST}_n(T_P,T_1,\ldots,T_n)$ instead of $P(T_1,\ldots,T_n)$. Now we can model the predicate $\mathbf{INST}_1$ on the real-world relation between objects and groups or sets of them (writing $T_1 \epsilon T_P$ instead of $\mathbf{INST}_1(T_P,T_1)$). This leads to the disposal of all other predicates by the well known way of coding n-tuples of terms and

[14] *Viz* Frege (1892b).

considering n-ary predicates as unary predicates of n-tuples (hence leaving the single predicate $\mathbf{INST}_1$, or, in the more usual notation, ϵ). Doing this, we can see the replacement of P's by T_P's as the replacement of ontologically problematic properties by unproblematic sets. This is what Russell (ibid., p. 51) calls "the principle of abstraction".[15]

The move just described can be considered as an example of "second-order regimentation". After carrying out the "first-order" regimentation to rid ourselves of the irrelevant idiosyncrasies of language, we went on to regiment the regimentation with the purpose of gaining a simpler, or in some sense more suitable, formal language. In this way we can reach formal languages whose relationship to natural language is far from straightforward.

If we consider these radically regimented languages, such as the language of set theory, we may be tempted to conclude that they amount to the true structure of reality, and to see natural language as reflecting this structure in a distorted way. If we take for granted that language depicts facts of the world (and to deny this seems to be to fly into the face of reason), then such a claim indeed is sound. However, we must realize how such a claim can be reasonably understood: saying that such simpler and more ordered languages reflect the world in a more direct way than the less simple and less ordered ones is in fact nothing more than a more dramatic way of saying that they do more accurately fit our standards of simplicity and order. They can be said to reflect the true structure of the world not because they would be based on an empirical investigation of the world which would reveal us its "true workings", but rather because they embody the "true structure" of language. More about all of this will follow in the subsequent chapters.

[15]Let us note that the operation of turning predicates into terms via the introduction of ϵ is straightforward only as long as we use $T \epsilon T_P$ as a mere short-cut for $P(T)$. In this way we reach only unproblematic, "virtual" sets; however, we must complicate the syntax of our schematization in an unwanted way (we have to secure that $T \epsilon T_P$ does, while $T_P \epsilon T$ does not, make sense). On the other hand, accepting ϵ as a fully-fledged predicate (so that $T_P \epsilon T$ is no less meaningful than $T \epsilon T_P$) we have to develop a theory of "real" sets including all its well-known perplexities. For a detailed discussion of the difference between "virtual" and "real" notions of a set see Quine (1969, esp. Part One).

2.7 Formal Languages

The role of all the symbolic means discussed so far is the role of auxiliaries which facilitate capturing consequence as it occurs in natural language. We have seen that systematic regimentation yields us a kind of symbolic language. However, such a language does not have any life of its own, it is a mere "shadow" of natural language. To speak about symbolic expressions is merely to use a convenient shorthand for speaking about their natural language counterparts; to say that it is valid to infer a symbolic statement from other symbolic statements is to say that this holds for their counterparts in natural language. Thus, if we are to justify the fact that it is valid to infer (3') from (2') and (1'), we say that (3') stands for (3), (2') for (2), and (1') for (1), and that (3) can be seen to be the consequence of (1) and (2); and if we are to justify the validity of the schema (15), we say that all its instances "can be seen" to be valid. Moreover, the language of symbolic expressions is essentially open: we are always prepared to extend and modify it to reflect any new aspects of natural language which may catch our attention.

However, we have stressed that logical theory aims at a canon, that it is to summarize valid inferences. To canonize is to bridge the gap between the descriptive and the normative, to take *what is the case* and to make *what should be the case* out of it. Hence, at some point we cease acknowledging that the validity of (15) depends upon its being a mere description of its natural-language instances, and we start to see it on its own as a *prescription*, as a *norm*; we begin to use the schema to *justify* the validity of the natural-language instances rather than the other way round. The justification of a descriptive pronouncement always lies in what it purports to describe. A normative pronouncement, on the other hand, is self-justifying: we have suspended the necessity to justify it externally.

This means that after reaching a certain stage of schematization we cut off the umbilical cord by which it is related to natural language and start to consider it as self-supporting. We articulate definite rules specifying what is a well-formed expression and what is not, and then we articulate rules to define what counts as valid inference (to be discussed in the following chapter). We turn the open family of symbolic expressions into a closed formal language. We no longer consider the validity of the inference of (3') from (2') and (1') to be a matter of the relationship of (1'), (2') and (3') to their natural-language counterparts. Instead, we start to take it simply as an unquestioned fact. However, this does not mean that we give up the possibility of questioning the norms altogether: what we give up is their *incessant* questioning. The adequacy now begins to be a matter of the language as a whole.

Let us consider the schema (1') that has arisen out of the regimentation of (1). If we consider it a matter of mere symbolization, then saying that it is false would be simply wrong - it would be to say that Aristotle is not human. On the other hand, if we consider (1') to be part of a formal language, then we may well say that it is false - because truth and falsity become matters internal to the formal language and nobody can keep us from defining just this kind of language. But then we have to say that such a language, if it purports to relate to natural language in such a way that (1') is to regiment (1), is not adequate.

So we have the disorderly natural language, on the one hand, and a formal, strictly rule-based language, on the other as its reconstruction. The reconstruction is to serve for normative, but also for "explicative" purposes. As Wittgenstein (1969, §36) puts it, "wollen wir, für unsere Zwecke, den Gebrauch eines Wortes bestimmten Regeln unterwerfen, so stellen wir seinem fluktuirenden Gebrauch einen anderen an die Seite, indem wir einen charakteristischen Aspekt des ersten in Regeln fassen."[16]

This is the step from the merely symbolic to the formal, or, as van Heijenoort (1967) calls it, from the notion of *logic as language* to the notion of *logic as calculus*, the step whose understanding is quite crucial for understanding modern logic and its problems. It was the confusion of these two views that underlay such crucial controversies as that between Frege and Hilbert, and it is this confusion that lays the foundation of contemporary misconceptions of the possibilities and limitations of logical theory.

The difference between a natural language and a formal one is that the latter is based on explicit rules, that it is *criterial*. We have a criterion that determines what is an expression of the language and what is not, i.e., we have a criterion of "well-formedness"; and we have a criterion that determines what is an instance of consequence. Note that we can never have any such criteria directly for natural language: we can put down standard orthographies, grammars and logics, but the ultimate decision will always be determined by the

---

[16]"If for our purposes we wish to regulate the use of a word by definite rules, then alongside its fluctuating use we set up a different use by codifying one of its characteristic aspects."

factual use of language, not by any standing criteria, however institutionally they may be guaranteed.

3 Axiomatic Systems

3.1 Necessary Truth and Logical Truth

Now we are ready to focus our attention on the main problem of logic, the capturing of consequence. Let us first point out the obvious fact that consequence is one side of the coin, the other side of which is necessary truth; and that we can equivalently talk of capturing necessary truth instead of capturing consequence. Each instance of consequence, if articulated, gives rise to a necessarily true statement: to say that it is valid to infer a statement S from statements $S_1,\ldots,S_n$ is to say that the statement *If (it is the case that) S_1 and ... and S_n, then (it is the case that) S*, or, regimented into symbolic notation, $(S_1 \wedge \ldots \wedge S_n) \rightarrow S$, is necessarily true. Thus the valid inference (16) can be expressed as the necessarily true statement (17); similarly, the validity of (18) is tantamount to the necessary truth of (19)[17].

$$\mathbf{Hu(Ar),\ \forall x\ Hu(x) \rightarrow Mo(x) \Rightarrow Mo(Ar)} \tag{16}$$
$$\mathbf{(Hu(Ar) \wedge (\forall x(Hu(x) \rightarrow Mo(x))) \rightarrow Mo(Ar)} \tag{17}$$
$$\mathbf{Hu(Ar) \Rightarrow Mo(Ar)} \tag{18}$$
$$\mathbf{Hu(Ar) \rightarrow Mo(Ar)} \tag{19}$$

On the other hand, every necessarily true statement can be understood as a special case of an instance of consequence, namely as the instance whose antecedent is void and whose consequent is the statement. This means that an account of consequence is *eo ipso* an account of necessary truth and vice versa; hence, if we have so far understood logic as the pursuit of consequence, we can now also, equivalently, consider it as the pursuit of necessary truth.

---

[17]The usual objection to reducing consequence to necessary truth is that the language we are considering may not contain the resources necessary for the reduction (equivalents of $\wedge$ and $\rightarrow$). However, we are not talking about formal languages, but about the *natural*, *pre-theoretical* notions of consequence and necessary truth; and we surely do have natural resources of articulating them (otherwise there could be no logic).

We have stressed that logical theory presupposes the existence of a pre-logical, "intuitive" consequence. In view of what has been stated we can now say that logic presupposes the existence of two pre-logical, "intuitive" oppositions: the opposition between truth and falsity, and that between necessity and contingency. However, let us stress that although the understanding of both these oppositions is assumed to be pre-theoretical or "intuitive", the understanding of each of them is quite a different matter. We understand the difference between truth and falsity without being always able to tell a true statement from a false one; to do so may be a matter of "investigation of the world". On the other hand, we understand the necessary/contingent opposition right off, for the ability to distinguish a necessary statement from a contingent one is a matter of knowing language[18]. To say that a statement is contingent is to say that it makes sense to think about it as true as well as false (although, of course, not as true and false simultaneously), that whether it is true or false, it just *happens* to be such. To say that a statement is necessary is to say that it makes no sense to think about it as changing its truth value; in many cases we may be ignorant of the truth value of a necessary statement, although we always know that it constantly possesses a definite one. The boundary between the contingent and the necessary then is also the boundary between two ways of finding the truth value of a statement in case we do not see it: to find a truth value of a contingent statement is the practical business of natural sciences, whereas to find that of a necessary statement is a theoretical armchair matter (although it may also be subject to systematic science - *viz* mathematics)[19].

So there is a distinction between necessary truths, which can be discovered "sitting in an armchair", and between contingent truths, the discovery of which is a matter of investigating the world and often is carried out by some kind of experts. However, a reflection of this distinction can be found within the realm of necessary truth itself. There are necessary truths

---

[18]At least to the extent to which the opposition makes sense at all. Thus far we are considering the opposition to be straightforward; ways to question this assumption will be dealt with in Chapter 10.

[19]Attempts are sometimes made to base answers to certain philosophical questions on distinguishing between the *a priori* and the analytic. (Kant is not alone in his well-known suggestion that there are statements that are *a priori*, but not analytic; Kripke, 1972, for example, argues that there are, the other way around, statements which are analytic, but not *a priori*.) However, we do not consider this kind of distinction useful for our purposes here.

which are limited to specific parts of our discourse and can thus in a sense be considered a matter of corresponding experts, of the authorities in the very part of discourse; and there are other necessary truths which are so general that they seem to support our overall discourse as such. The necessary truth of a statement such as *Whales are mammals* can be considered to be restricted to zoologists; to be ignorant of it does not necessarily mean not to know English. On the other hand, to be aware of the necessary truth of a statement like *If whales are mammals, then whales are mammals* is *conditio sine qua non* of knowledge of language. The necessary truth of (19) may be considered as being of the former kind, that of (17) is of the latter kind. If someone doubted the truth of (19), he might be referred to an anthropologist for explanation[20]; should he, on the other hand, doubt (17), it would be clear that he simply does not understand English.

We have said that logic is to account for consequence or necessary truth; but its direct concern is only the most general, universal part of consequence and necessary truth, which then is usually called *logical consequence* or *logical truth*. (16) is considered an instance of logical consequence and (17) a logical truth; whereas (18) and (19) are considered as a matter extraneous to logic.

So we can distinguish the following three pre-theoretical "modes of truth":

TRUTH ⇐ NECESSARY TRUTH ⇐ LOGICAL TRUTH

Table 2

In the previous chapter we decided to call a schema *valid* if all its instances are true statements. Now, having introduced new, stronger "modes of truth", namely necessary truth and logical truth, we can analogously introduce stronger notions of validity, namely *necessary validity* and *logical validity*. A schema is called valid iff all its instances are true; it will be called necessarily

[20]In fact, this is not a very good example - everybody can in a sense count as an expert with respect to what it is to be human; and there are probably people who would deny the necessary truth of (17), although they could be hardly considered incompetent as speakers of English.

valid iff all its instances are necessarily true, and it will be called logically valid iff all its instances are logically true[21]. Thus, the schemata (20) and (21), resulting from (17) and (19) via abstraction, are not only valid, but also necessarily valid; (20) is, moreover, logically valid.

$$(\boldsymbol{P}_1(\boldsymbol{T}) \wedge (\forall x(\boldsymbol{P}_1(x) \rightarrow \boldsymbol{P}_2(x))) \rightarrow \boldsymbol{P}_2(\boldsymbol{T}) \tag{20}$$

$$\mathbf{Hu}(\boldsymbol{T}) \rightarrow \mathbf{Mo}(\boldsymbol{T}) \tag{21}$$

Every necessarily valid schema is surely valid; so necessary validity is a stronger concept than validity *simpliciter*. However, is it nontrivially stronger? Is there a schema that is valid, but not necessarily valid? The answer to this question is important; but it depends on the way we handle the term *true*. In fact, there are two quite different ways, and they lead to two different concepts of validity. To illustrate this let us consider the schema (22), where **Fe** schematizes the predicate *featherless* and **Bi** the predicate *biped*.

$$(\mathbf{Fe}(\boldsymbol{T}) \wedge \mathbf{Bi}(\boldsymbol{T})) \rightarrow \mathbf{Hu}(\boldsymbol{T}) \tag{22}$$

Is this schema valid or not? Does it have a false instance, i.e is there some non-human featherless biped? Let us consider the statement (23), where **Fr** schematizes *Frodo Baggins*, the name of the famous hero of J.R.R.Tolkien's book *Lord of the Rings*, who is a hobbit, a dwarf-like creature walking on two feet, but not being human.

$$(\mathbf{Fe}(\mathbf{Fr}) \wedge \mathbf{Bi}(\mathbf{Fr})) \rightarrow \mathbf{Hu}(\mathbf{Fr}) \tag{23}$$

Can we say, that (23) is a false instance of (22)? If so, then (22) would clearly fail to be valid. In one sense, (23) clearly *is* a false instance of (22), thereby disproving its validity: Frodo is a featherless biped, but he is not human. However, we can also insist that there is "in fact" no Frodo Baggins, hence that **Fr** stands for nothing, and that hence (23) is not true, but meaningless; so in another sense, (23) is not a false instance of (22).

Two possible answers to the question exemplify two essentially different notions of truth and consequently of validity. One of them, the one

[21]Note that there is no distinction between validity and necessary validity in the case of inferences: what we have called *valid inference* is essentially *necessarily valid inference*!

disqualifying (23) as an instance to (22) on the basis of its fictional status, amounts to *truth* in its primordial sense, to *actual truth within our real world*, the other one rather amounts to *any imaginable truth*, or *truth in any imaginable world*. It is surely the first concept that is basic and invites our adherence; however, there are two independent reasons why this first concept is not suitable for our present purpose.

First, this view of truth presupposes a straightforward and clear-cut dividing line between "reality" and "imagination"; which is notoriously problematic[22]. Second, if we based validity upon this concept of truth, then determining validity would be a matter of investigating the world, and hence of natural science, not of logic. The other concept of validity is independent from the actual status of the world and hence is a fair game for logic.

This leads us to basing the concept of validity on truth in the other sense, in the sense in which each intelligible statement counts as an instance, irrespectively of whether it is a part of a scientific theory or a fiction. Taking validity in this sense, it is far from clear that necessary validity is something more than validity *simpliciter*: it is hard to find an example to a schema that would be valid but not necessarily valid[23].

---

[22]It is not only that in many cases we cannot tell reality from imagination; but sometimes such telling simply makes no sense. Consider, e.g., *thoughts* - a behaviourist psychologist would say that they are the product of the imagination of his non-behaviourist colleagues; but his colleagues would swear that thoughts are utmostly real. Is then *Thoughts are not material* a true instance of *x is not material*, or not?

[23]However, if we considered languages defined formally, then there, of course, would not be a problem with finding such an example. Take, for instance, the formal language which would result from freezing our regimentation in the stage where it contains only the constant terms **Ar** and **Pl** and the constant predicates **Hu** and **Mo** (and letting the statements retain their "natural" characteristics, so that all of **Hu(Ar)**, **Hu(Pl)**, **Mo(Ar)**, **Mo(Pl)** are true, but none of them is necessarily true). Then we have *P*(*T*) as a schema that is valid (all its instances are true), but not necessarily valid (not every of its instances - in fact none of them - is *necessarily* valid).

3.2 Truth and Validity

The concept of validity (and the concepts of necessary and logical validity) apply to schemata; but they can also, derivatively, be applied to statements. A statement can be called valid if it is an instance of a valid schema, and similarly for necessary and logical validity. However, we must keep in mind that in contrast to the concepts of necessary and logical truth, which are assumed to be grounded in our pre-theoretical "intuition" (we "intuitively" know what it is to be necessarily and logically true before we start to do logical theory), the concepts of validity, of necessary validity, and of logical validity are outcomes of the theory. So whereas the former concepts are what is to be accounted for by the theory, the latter ones are what is brought to life by the theory in order to be utilized as means of such an account.

It is obvious that a statement that is valid is true, and similarly for necessary and logical validity and truth. So we have the following dependencies

$$\begin{array}{ccc}
\text{TRUTH} & \Leftarrow & \text{VALIDITY} \\
\Uparrow & & \Uparrow \\
\text{NECESSARY TRUTH} & \Leftarrow & \text{NECESSARY VALIDITY} \\
\Uparrow & & \Uparrow \\
\text{LOGICAL TRUTH} & \Leftarrow & \text{LOGICAL VALIDITY}
\end{array}$$

Table 3

The concept of validity resulted from our recognition of the fact that some instances of consequence, i.e. some necessarily true statements, are insensitive to the variation of certain parts; that expressions of certain categories can be freely interchanged within them without violating their necessary truth. Of the categories of expressions we introduced in the course of our discussion of the ways of schematization of language, terms - with some possible exceptions, like the terms of mathematics - appear to exhibit such indifference with respect to all necessary truths. Predicates exhibit it with respect to some of them, and certainly with respect to all *logical* truths - though again there are some possible exceptions such as the predicate of identity.[24] Thus, in (17) we can vary both the term and the predicate; however, in (19) we can only vary the term freely, not the predicates.

The fact that the constants of the mentioned categories can be interchanged within logical truths without limitation can be understood to mean that these expressions are not directly "responsible" for logical truth. So it appears plausible to call the constants of the former kind *extralogical*, and those of the other kind, those which are usually irreplaceable, *logical*. Hence predicate and term constants (with the mentioned possible exceptions) are extralogical, whereas operator and quantifier constants are logical.

We have introduced parameters to "abstract away" extralogical expressions; and in this way we have reached schemata and the concept of validity. Let us consider two basic kinds of schemata: the first of them resulting from "abstracting away" schema's term(s) (**Hu**(***T***), **Ad**($\boldsymbol{T_1}$,$\boldsymbol{T_2}$), etc.), the second from "abstracting away" both its term(s) and its predicate(s) (***P***(***T***), ***P***($\boldsymbol{T_1}$,$\boldsymbol{T_2}$), etc.)[25]. These two levels of abstraction can be followed by further levels resulting from abstracting away statements; but these are not important for us now. Note that some of the "first-level" schemata will be identical with statements: the "first-level" schema of a statement which contains no terms (such as ∀x**Hu**(x)) does not differ from the statement itself.

Returning to our Table 3: to simplify it, let us consider the following fact. A schema results from its instances by "abstracting away" some of their extralogical parts, different instances of the same schema can differ one from another at most in extralogical constants. This means that if an instance of a schema is logically true, every other instance of the same schema must be logically true, too. Hence if a statement is logically true, it is *eo ipso* logically valid, and this clearly means that the concepts of logical truth and logical validity coincide.

Moreover, the whole idea of logical schematization was driven by the tacit assumption that necessary truth is capable of being captured in a

---

[24]Some logical truths are, further, insensitive to the variation of statements. These are the "most logical" of the logical truths; they are the truths captured by the propositional calculus.

[25]Other kinds of schemata, such as **Ad**(**Pl**,***T***) (i.e. with only some of their terms "abstracted away"), or ***P***(**Ar**) (i.e. with predicates, but not terms "abstracted away"), are not interesting for us.

schematic, abstract manner. This assumption amounts to the idea that necessarily true statements are instances of necessarily valid schemata and are thus themselves necessarily valid; or, in other words, that out of the ways we can schematize a necessarily true statement at least one leads us to a necessarily valid schema. However, if there is a necessarily valid schema instantiated by a statement, then that statement is - by definition - necessarily valid; hence the concepts of necessary truth and necessary validity coincide.

These conclusions allow us to simplify Table 3 to Table 4.

TRUTH	⇐	VALIDITY	⇐	NECESSARY TRUTH	⇐	LOGICAL TRUTH
				=		=
				NECESSARY VALIDITY		LOGICAL VALIDITY

Table 4

The reductions leading us from Table 3 to Table 4 were based on assumptions essential to our very framework. But we can try to make more eccentric conjectures to facilitate further reductions. Understanding truth and validity in the broad sense (allowing for Frodo Baggins to be taken as an example of a featherless non-human biped), we can conjecture that every valid schema is also necessarily valid; which implies the identification of validity with necessary validity and hence also with necessary truth. Let us call this the *first validity conjecture*: *If a schema is valid, then it is necessarily valid, i.e. its instances are not only true, but necessarily true*. Equivalently: a schema that has an instance that is not necessarily true has also an instance that is not true[26].

Moreover, besides conjecturing the identification of necessary truth with validity, we can also conjecture the identification of logical truth with a notion of more general validity. We can make sense of such a notion in terms of the above introduced distinction between first-level and second-level schemata. The second-level schemata contain only that which is relevant from the point of view of logical truths and thus will be called *general schemata* (or *logical forms*);

[26]This amounts to the idea which underlies the introduction of possible worlds into semantics: that if something *can be* false (it is not necessarily true), then it *is* false in some would-be, imaginary context or world.

and a statement will be called *generally valid* if it is an instance of a valid general schema. Now we can conjecture that a valid general schema is logically valid; which implies that general validity coincides with logical validity and hence with logical truth. Let us call this the *second validity conjecture*: *If a general schema is valid, then it is logically valid, i.e. its instances are not only true, but logically true.* Equivalently: a general schema that has an instance that is not logically true has also an instance that is not true.

These validity conjectures make it possible to reduce the four "modes of truth" of Table 4 to the three of Table 5; notice that all the three remaining modes are based on the single pre-theoretical concept of truth *simpliciter*.

TRUTH	⇐	VALIDITY	⇐	GENERAL VALIDITY
		=		=
		NECESSARY TRUTH		LOGICAL TRUTH

Table 5

The validity conjectures articulate our intuition that necessary truths are in a sense general, universal truths, and that logical truths are the most general of the universal truths. However, we leave them as conjectures, without subscribing to them.

3.3 An algebraic perspective

It may be of helpful to reiterate all what has been said so far in a simple algebraic setting. However, it is absolutely crucial to understand the way we employ the algebraic apparatus: our aim is not to set up a mathematical theory, but only to build a schematic picture to facilitate the understanding of the facts which were stated informally before. The reason for meticulously remembering that our algebraic articulation is merely an illustrative picture is that were we to take it literally we would lose the vital contrast between the pre-formal and the formal: to do the algebraization we have to express even the pre-formal formally.

What we consider as (virtually) given prior to logical theory may be considered as a class S (of statements), its subclass T (of true statements), the subclass NT (of necessary true statements) of T, and the subclass LT (of logically true statements) of NT. We assume that our logico-grammatical analysis yields the class Σ (of schemata) together with a mapping INST of Σ into the power set of S ($\mathrm{INST}(\sigma)$ yielding the class of instances of the schema σ), and a subclass Σ^* (of general, or "second-level", schemata) of Σ.

We assume that schemata are chosen in such a way (via stipulating which constants are extralogical and hence replaceable by constants) that each statement is an instance of a general schema and that passing from an instance of a schema to another instance of the same schema cannot take us out of the realm of logical truths; hence we assume (F1) and (F2). Moreover, we assume that each necessary truth can be captured schematically, that every necessarily true statement is an instance of a schema whose instances are all necessarily true; hence we assume (F3).

If $s \epsilon \mathrm{S}$, then there is an $\sigma \epsilon \Sigma^*$ such that $s \epsilon \mathrm{INST}(\sigma)$ (F1)

If $\sigma \epsilon \Sigma$ and $\mathrm{INST}(\sigma) \cap \mathrm{LT} \neq \varnothing$, then $\mathrm{INST}(\sigma) \subseteq \mathrm{LT}$ (F2)

If $s \epsilon \mathrm{NT}$, then there is a $\sigma \epsilon \Sigma$ such that $s \epsilon \mathrm{INST}(\sigma)$
and $\mathrm{INST}(\sigma) \subseteq \mathrm{NT}$ (F3)

We define the subclasses V^Σ (of valid schemata), NV^Σ (of necessarily valid schemata), and LV^Σ (of logically valid schemata) of Σ as follows

If $\sigma \epsilon \Sigma$, then $\sigma \epsilon \mathrm{V}^\Sigma$ iff $\mathrm{INST}(\sigma) \subseteq \mathrm{T}$ (D1)

If $\sigma \epsilon \Sigma$, then $\sigma \epsilon \mathrm{NV}^\Sigma$ iff $\mathrm{INST}(\sigma) \subseteq \mathrm{NT}$ (D2)

If $\sigma \epsilon \Sigma$, then $\sigma \epsilon \mathrm{LV}^\Sigma$ iff $\mathrm{INST}(\sigma) \subseteq \mathrm{LT}$ (D3)

It is now obvious that

$$\mathrm{LV}^\Sigma \subseteq \mathrm{NV}^\Sigma \subseteq \mathrm{V}^\Sigma \qquad \text{(F4)}$$

We can further define the subclasses V (of valid statements), NV (of necessarily valid statements), and LV (of logically valid statements) of S as follows

If $s \epsilon S$, then $s \epsilon V$ iff there is a $\sigma \epsilon \Sigma$ such that $s \epsilon INST(\sigma)$ and $\sigma \epsilon V^{\Sigma}$. (D4)

If $s \epsilon S$, then $s \epsilon NV$ iff there is a $\sigma \epsilon \Sigma$ such that $s \epsilon INST(\sigma)$ and $\sigma \epsilon NV^{\Sigma}$. (D5)

If $s \epsilon S$, then $s \epsilon LV$ iff there is a $\sigma \epsilon \Sigma$ such that $s \epsilon INST(\sigma)$ and $\sigma \epsilon LV^{\Sigma}$. (D6)

(F4) now implies (F5); and (F6) through (F8) are evident.

$$LV \subseteq NV \subseteq V \qquad (F5)$$
$$V \subseteq T \qquad (F6)$$
$$NV \subseteq NT \qquad (F7)$$
$$LV \subseteq LT \qquad (F8)$$

This gives all the relationships summarized in Table 3 of the previous paragraph.

Now let us assume that $s \epsilon LT$. Then, according to (F1), there is a $\sigma \epsilon \Sigma$ such that $s \epsilon INST(\sigma)$; and, according to (F2), $INST(\sigma) \subseteq LT$. This, according to (D3), means that $\sigma \epsilon LV^{\Sigma}$, and this, according to (D6), in turn means that $s \epsilon LV$. Hence we have proven (F9), which together with (F8) gives (F10).

$$LT \subseteq LV \qquad (F9)$$
$$LT = LV \qquad (F10)$$

Now let us assume that $s \epsilon NT$. Then, according to (F3), there is a $\sigma \epsilon NV^{\Sigma}$ such that $s \epsilon INST(\sigma)$; hence $s \epsilon NV$. This gives (F11) and together with (F7) it further gives (F12); whereby we have reached the relationships of Table 4.

$$NT \subseteq NV \qquad (F11)$$
$$NT = NV \qquad (F12)$$

What we called validity conjectures can now be formulated as follows:

$$V^{\Sigma} \subseteq NV^{\Sigma} \qquad (VC1)$$
$$V^{\Sigma} \cap \Sigma^{*} \subseteq LV^{\Sigma} \qquad (VC2)$$

Given (F4), (VC1) implies the identity of V^{Σ} and NV^{Σ}, and hence, in view of (D4) and (D5), the identity of V and NV. Given (F12), this yields (VC1').

$$NT=V \tag{VC1'}$$

Now let us define the class of generally valid statements by (D7); (F13) then is obvious.

$$\text{If } s\in S\text{, then } s\in GV \text{ iff there is an } \sigma\in\Sigma^{*} \text{ such that } s\in INST(\sigma) \text{ and } \sigma\in V^{\Sigma}. \tag{D7}$$

$$GV\subseteq V \tag{F13}$$

Let us assume that $s\in LT$. According to (F1), there is a $\sigma\in\Sigma^{*}$ such that $s\in INST(\sigma)$; and, according to (F2), $INST(\sigma)\subseteq LT$. As $LT\subseteq T$, this implies that $\sigma\in V^{\Sigma}$; and this, according to (D7), means that $s\in GV$. This establishes (F14).

$$LT\subseteq GV \tag{F14}$$

(VC2) obviously implies that if $s\in GV$, then $s\in LV$; hence given (F10), that $GV\subseteq LT$. Hence given (F14), (VC2) implies (VC2'). Thus we have the relationships of Table 5.

$$LT=GV \tag{VC2'}$$

Notice that whereas we could not use (F12) to *define* NT (because the definition would be circular - NV is defined via NV^{Σ} and NV^{Σ} is in turn defined via NT), we *could* use (VC1') to define NT (for GV is *not* defined using NT). Similarly we could not use (F10) to define LT, but we could use (VC2'). This fact underlies the claim that whereas it makes no sense to reduce necessary truth to necessary validity (or logical truth to logical validity), it does make sense to reduce necessary truth to simple validity (or logical truth to general validity). The latter reductions, however, presuppose the acceptance of the validity conjectures.

3.4 Axiomatic Systems and Formal Calculi

Abstractive schematization allows us to represent an infinite number of statements by a single schema, which in turn allows us to study valid schemata instead of true statements. As a schema may cover an infinite number of instances, this allows us to address an infinite range of concrete statements via addressing a finite number of schemata. However, even schemata which summarize logical truths prove to be infinite in number; so we are unable to capture *all* the logically true statements, let alone all necessarily true statements, by means of a finite number of schemata[27]. This is to say that the concepts of validity do not lead directly to a criterial reconstruction of the concept of logical truth.

Thus, allthouth we no longer need to summarize all the logically valid statements, but only all the logically valid schemata, we still cannot carry this out by simply making a list. We can, however, make a "potential list"; we can give a prescription for making (potentially) an infinite list. We can formulate an *axiomatic system*: we can give a finite list of basic logically valid schemata (*axiom schemata*) and furnish rules for passing from logically valid schemata to further logically valid schemata (*inference rules*). A *proof* in such a system is a sequence of schemata in which each of the elements is either an axiom or follows from some previous elements of the sequence according to an inference rule. A schema is called *provable* if there exists a proof which ends with it. A statement is a *theorem*, if it is an instance of a provable schema[28]. We can, of course, equally well speak about directly proving statements (rather than schemata); we can take the instances of axiom schemata as *axioms* and look at

---

[27]This can be seen from the fact that such a sequence as $P(T) \rightarrow P(T)$, $P(T) \wedge P(T) \rightarrow P(T)$, $P(T) \wedge P(T) \wedge P(T) \rightarrow P(T)$, ... contains only valid schemata and it can be continued without limitation.

[28]At the beginning of the previous chapter we considered a theorem as an inference of a V' from a V and its proof a decomposition of the inference $V \Rightarrow V'$ into a chain $V = V_1 \Rightarrow V_2 \Rightarrow \ldots \Rightarrow V_{n-1} \Rightarrow V_n = V'$ of "simple" inferences. The present picture, resulting from our switching from consequence to necessary truth, leads us instead to say that what we in this case prove is the necessary truth $V \rightarrow V'$, and that we do it with the help of the "simple", "evident" necessary truths $V_1 \rightarrow V_2, \ldots, V_{n-1} \rightarrow V_n$.

the schematic inference rules as summarizations of rules for passing from statements to statements.[29].

A formal language plus an axiomatic system is called a *formal calculus*; if the theorems of the system are to be understood as regimenting logical truths, then we speak about a *logical* calculus. In case of a logical calculus, the pre-theoretical property *to be logically true* is reconstructed as the theoretical *to be provable*. This is to say that the criterion of logical truth arrived at by axiomatic method is *provability*[30]. Let us therefore call this kind of account of logical (or, more generally, necessary) truth *proof-theoretic*[31].

However, it is again crucial to realize the essential distinction between an axiomatic system meant as a mere symbolic capturing of natural necessary truth, and a *formal* axiomatic system. The former is descriptive, it merely helps us summarize some necessary truths as appearing in natural language; the latter is, on the other hand, stipulated and self-contained, and it can be used for prescriptive purposes. The misunderstanding which may arise from not keeping these two notions properly apart is illustrated by the controversy between Frege and Hilbert. In a letter to Frege Hilbert writes[32]:

> You write: 'From the truth of the axioms it follows that they do not contradict one another'. I was very interested to read this particular sentence of yours, because for my part, ever since I have been thinking, writing and lecturing about such matters, I have

---

[29]We can, of course, avoid schemata altogether and employ substitution as an inference rule. This makes no substantial difference - schemata and substitution are clearly two sides of the same coin. See later Section 4.6.

[30]Note, however, that even if we were able to characterize the class of all the valid inferences, or all the valid statements, in this way, the characterization would be a *criterion* only in a weak sense of the word. It would not necessarily be a criterion in the sense that it would always allow us to determine whether a given inference is valid or not. See further Section 4.10.

[31]Formal calculi may be intended to summarize not only logical truths, but some more extensive range of necessary truths. Only logical calculi are wholly within the competence of a logician; other calculi require that logicians join forces with experts in the domain corresponding to the necessary truths summarized, the former contributing inference rules and logical axioms, the latter delivering substantial axioms.

[32]See Steck (1941).

> been accustomed to say just the reverse: if the arbitrarily posited axioms are not in mutual contradiction with the totality of their consequences, then they are true - the things defined by the axioms exist.

The clash is evidently the one between logic as a symbolism (van Heijenoort's *logic as language*) and as a formalism (his *logic as calculus*). For Frege, a logical formula is a mere regimentation of a natural language statement, and to say that it is an axiom is to say that the statement for which it stands is evidently (necessarily or logically) true - hence assuming that our natural language is consistent (what could it mean for natural language to be inconsistent?) we may conclude that no two formulas acknowledged as axioms can be in contradiction. For Hilbert, on the other hand, logical formulas are elements of an abstract, self-contained system; and the system is reasonable (in which case its axioms can be considered true) if it fulfils certain formal criteria, especially if it is not contradictory.

3.5 Material Soundness and Completeness

If we have a logical calculus, i.e. a formal language plus an axiomatic system, we can consider whether its formulas can be put into correspondence with statements of natural language in such a way that its theorems correspond to logically true statements. To posit such a correspondence is to carry out logical analysis of language; and to consider whether theorems do or do not correspond to logical truths is to consider the adequacy of the analysis. The calculus can be said to be *(materially) sound* if all its theorems correspond to logical truths; and it can be called *(materially) complete*, if to every logical truth there corresponds a theorem. To say that a logical calculus is sound and complete is to say that the class of its theorems conicides with the class of logical truths; it is hence to say that the calculus yields a precise reconstruction of logical truth. The same holds for formal calculi aiming at capturing broader ranges of necessary truths.

It is essential to realize that soundness and completeness in this sense are concepts utterly different from the more usual concepts of formal (model-theoretical) soundness and completeness to be discussed in the next chapter. In particular, material soundness and completeness can never be subject to a formal proof. Formal proof is always a matter internal to a formal reconstruction, and therefore cannot prove anything *about* the reconstruction (as a whole).

This is also the point at which the limitations of our algebraic depiction presented above can be clearly seen. To accommodate the concepts of material soundness and completeness into our picture would involve viewing axioms as a finite subset A of the set S and inference rules as functions from the power set of S to S, and defining the class TH of theorems as the smallest set containing A and closed with respect to inference rules. The problem of soundness and completeness of the system would thus become the problem of the coincidence of TH with LT (or, in the more general case, with some more extensive part of NT). However, this would suggest the idea that it is a problem to be formally solved by algebraic (or set-theoretic) means; and this would be essentially mistaken. The step quite crucial for logical theory is the step from the pre-formal ("natural", material) to the formal, and if we try to depict this within the framework of a formal system, we let it appear as a trivial move internal to the system.

Let us further notice that the requirements of soundness and completeness are of different characters. Soundness seems to be a *conditio sine qua non* of a reasonable logical calculus, and we can also conceive of its proof (although not of a formal proof). To ascertain a logical calculus as sound it is enough to ascertain that its axioms are logically true and that its inference rules preserve logical truth.

In contrast to this, completeness is far more problematic both in respect of being required and of being proven. We may want to summarize only some particular kind of logical truths, and thus we could appreciate a system which is not complete in an absolute sense. In the case of logical truths, the extent of completeness we may reasonably require may be seen as a matter of the depth of the logical analysis we are actually performing, and in the case of more general classes of necessary truths it is, moreover, a matter of the domain of discourse we are investigating. Besides this, both logical and necessary truths amount to classes of statements which are notoriously fuzzy.

To sum up: the starting-point of logic is the class of (intuitively) valid inferences or (intuitively) necessarily true statements. A formal calculus provides a criterial, rule-based specification of a new class which is to coincide with (a part of) the original one (in the case of logical calculus with the class of logical truths); if the latter really can be considered as coinciding with the former, i.e. if the logical analysis yielded by comparing the formal calculus with language can be considered adequate, then the class specified via the axiomatics can be considered as a *reconstruction* of the original class of necessary truths. The reconstruction is "better" than the reconstructed class in that it is based on explicit rules, i.e. in that it is connected with an explicit criterion. To replace the original class with the new one is to make it more transparent and comprehensible.

3.6 Material Interpretation

The problem of adequacy of a logical calculus is, as we have seen, the problem of whether being a theorem coincides with being logically true, whether all instances of the schemata provable in the calculus are logically true, i.e. whether all the provable schemata are logically valid. Now to say that a schema is logically valid is to say that it is turned into a logically true statement by every conceivable replacement of parameters by constants. If we call the mapping of parameters onto constants *material interpretation*, then every such interpretation induces a mapping of schemata on statements. We can say that an interpretation *verifies (necessarily verifies, logically verifies)* a schema or a group of schemata if it maps the schema on a true (necessarily true, logically true) statement; and we can say further that an interpretation is a *material model (necessary material model, logical material model)* of a schema or a group of schemata if it verifies (necessarily verifies, logically verifies) the schema or each schema of the group. If we identify ***T*** with **Ar**, the schema **Hu(*T*)** yields the true (but not necessarily true) statement **Hu(Ar)**; hence every interpretation which maps ***T*** on **Ar** verifies **Hu(*T*)** (but does not necessarily verify it).

Interpretation in this sense is an inverse of abstraction, it is a "dis-abstraction".[33] Abstraction means the replacement of constants by parameters to obtain abstract schemata from concrete statements; interpretation means the opposite, the replacement of parameters by constants to obtain concrete statements from abstract schemata. Abstraction is many-to-one (it

[33]It is what Goodman (1978) calls *exemplification*.

assigns a single schema to a class of statements)[34] and interpretation is one-to-one; hence what corresponds to the inversion of abstraction is a whole space of interpretations.

This terminology allows us to express the concept of validity (necessary validity, logical validity) in a new way: a schema is valid (necessarily valid, logically valid) iff it is verified (necessarily verified, logically verified) by every interpretation, iff every interpretation is its model (necessary model, logical model) . Hence a statement is valid (necessarily valid, logically valid) iff it is an instance of a schema which is verified (necessarily verified, logically verified) by every interpretation.

Let us return to our algebraic picture. We assume that we have the class INT (of material interpretations) of functions from Σ to S such that the following holds:

If $s \epsilon S$ and $\sigma \epsilon \Sigma$, then $s \epsilon INST(\sigma)$ if and only if $i(\sigma)=s$ for some $i \epsilon INT$ (D8)

Now we can define the functions MOD, NMOD, LMOD mapping classes of schemata into classes of interpretations as follows:

If $\Sigma' \subseteq \Sigma$ and $i \epsilon INT$, then $i \epsilon MOD(\Sigma')$ iff $i(\sigma) \epsilon T$ for every $\sigma \epsilon \Sigma'$. (D9)

If $\Sigma' \subseteq \Sigma$ and $i \epsilon INT$, then $i \epsilon NMOD(\Sigma')$ iff $i(\sigma) \epsilon NT$ for every $\sigma \epsilon \Sigma'$. (D10)

If $\Sigma' \subseteq \Sigma$ and $i \epsilon INT$, then $i \epsilon LMOD(\Sigma')$ iff $i(\sigma) \epsilon LT$ for every $\sigma \epsilon \Sigma'$. (D11)

It obviously follows from (D8) that for every schema σ, $INST(\sigma)$ coincides with the class $\{s \mid i(\sigma)=s \text{ for some } i \epsilon INT\}$. This implies that $INST(\sigma) \subseteq T$ if and only if $i(\sigma) \epsilon T$ for every $i \epsilon INT$. In view of (D1), this further implies that $\sigma \epsilon V^{\Sigma}$ iff $i(\sigma) \epsilon T$ for every $i \epsilon INT$; and this, according to (D9), means that $\sigma \epsilon V^{\Sigma}$ iff $i \epsilon MOD(\{\sigma\})$ for every $i \epsilon INT$. Hence we have proved (F15); we can prove (F16) and (F17) analogously.

[34]Of course if we keep the range of constants to be obligatorily replaced by parameters fixed. Otherwise, more than one schema could result from every single statement containing at least one parameter.

$\sigma \epsilon V^{\Sigma}$ iff $MOD(\{\sigma\})=INT$ (F15)
$\sigma \epsilon NV^{\Sigma}$ iff $NMOD(\{\sigma\})=INT$ (F16)
$\sigma \epsilon LV^{\Sigma}$ iff $LMOD(\{\sigma\})=INT$ (F17)

(F16) together with (D5) and (VC1') then give (VC1''); whereas (F17) plus (D6) plus (VC2') yield (VC2'').

If $s \epsilon S$, then $s \epsilon NT$ iff there is an $\sigma \epsilon \Sigma$ such that $s \epsilon INST(\sigma)$ and $MOD(\{\sigma\})=INT$. (VC1'')
If $s \epsilon S$, then $s \epsilon LT$ iff there is an $\sigma \epsilon \Sigma^*$ such that $s \epsilon INST(\sigma)$ and $MOD(\{\sigma\})=INT$. (VC2'')

We have stated that a logical calculus related to language is adequate, or sound and complete, if all and only schemata which can be proved in it are logically valid. Now, after having introduced the concepts of material interpretation and model, we conclude that a schema is logically valid if and only if it is logically verified by every interpretation; hence we can now say that a logical calculus is materially sound if all its theorem-schemata are logically verified by every material interpretation; and that it is materially complete if only its theorem-schemata are logically verified by every material interpretation. A more general formal calculus is then materially sound (materially complete) if all (only) its theorem-schemata are necessarily verified by every material interpretation.

If we accept the validity conjectures, then we can go further and conclude that a formal calculus is materially sound if all its theorem-schemata are (simply) verified by every material interpretation; and that it is materially complete if only its theorem-schemata are (simply) verified by every material interpretation.

3.7 Interpretation vs. Truth-valuation

An interpretation maps schemata on statements, and each statement is either true, or false. Thus, every interpretation induces a *truth-valuation*, a mapping of schemata on truth-values; to say that a schema is verified by an interpretation is to say that it is mapped on *truth* by the truth-valuation induced by it.

Algebraically

If f is a function from Σ to S, then f^* is the function from Σ to $\{0,1\}$ such that $f^*(s)=1$ iff $f(s)\epsilon T$ (D12)

$INT^* = \{f^* \mid f\epsilon INT\}$ (D13)

Now to define the concept of validity we do not require the class INT, it is enough to have INT^*; i.e. we do not need interpretations, we need only corresponding truth-valuations. The point is that we can define MOD^* as in (D14) and then we evidently have (F18), (F19) and (F20).

If $\Sigma' \subseteq \Sigma$ and $f^*\epsilon INT^*$, then $f^*\epsilon MOD^*(\Sigma')$ iff $f^*(\sigma)=1$ for every $\sigma\epsilon\Sigma'$. (D14)

$\sigma\epsilon V^{\Sigma}$ iff $MOD^*(\{\sigma\})=INT^*$ (F18)

$s\epsilon V$ iff there is a $\sigma\epsilon\Sigma$ such that $s\epsilon INST(\sigma)$ and $MOD^*(\{\sigma\})=INT^*$ (F19)

$s\epsilon GV$ iff there is a $\sigma\epsilon\Sigma^*$ such that $s\epsilon INST(\sigma)$ and $MOD^*(\{\sigma\})=INT^*$ (F20)

Hence a statement is valid iff it is verified by every truth-valuation from INT^*. Now if we accept our *validity conjectures*, we can conclude that a statement is necessarily true iff it is an instance of a schema verified by every truth-valuation from INT^*, and that it is logically true iff it is an instance of a general schema verified by every truth-valuation from INT^*. In other words, we have

$s\epsilon NT$ iff there is a $\sigma\epsilon\Sigma$ such that $s\epsilon INST(\sigma)$ and $MOD^*(\{\sigma\})=INT^*$ (VC1''')

$s\epsilon LT$ iff there is a $\sigma\epsilon\Sigma^*$ such that $s\epsilon INST(\sigma)$ and $MOD^*(\{\sigma\})=INT^*$ (VC2''')

This fact may suggest that it might be possible to account for logical truth wholly bypassing the detour of assigning constants to parameters and instead directly considering assignments of truth values to schemata. This might seem to cut us off from the weird intricacies of language and to reformulate the whole matter of material soundness and completeness as a formal problem after all: a logical calculus is sound and complete iff every provable schema is verified by every admissible truth-valuation, and if every schema that is so verified is provable.

However, this is a mere illusion. The snag is in the word *admissible*: the above definition does not work unless we know which truth-valuations are admissible (which belong to INT*); and this is not possible without taking recourse to language. Let us consider the following schemata.

$$\boldsymbol{P_1}(\boldsymbol{T}) \tag{24}$$

$$\boldsymbol{P_2}(\boldsymbol{T}) \tag{25}$$

$$\forall x(\boldsymbol{P_1}(x) \rightarrow \boldsymbol{P_2}(x)) \tag{26}$$

Mapping the parameters $\boldsymbol{P_1}$, $\boldsymbol{P_2}$, $\boldsymbol{T}$ on constants we arrive at various interpretations of the schemata and thereby at various truth-valuations of them. Thus mapping $\boldsymbol{P_1}$ on **Hu**, $\boldsymbol{P_2}$ on **Mo** and $\boldsymbol{T}$ on **Ar** we arrive at the truth-valuation that maps all the three schemata on *truth*. Mapping $\boldsymbol{P_1}$ on **Mo**, $\boldsymbol{P_2}$ on **Hu** and $\boldsymbol{T}$ on **Ar** leads to the truth-valuation which maps (24) and (25) on *truth*, but (26) on *falsity*. If we use constants corresponding to various expressions of natural language, then we soon reach all possible distributions of truth values among the three schemata with one exception - no mapping of $\boldsymbol{P_1}$, $\boldsymbol{P_2}$, $\boldsymbol{T}$ on constants would lead to an interpretation verifying (24) and (26) and falsifying (25). This means that the space of admissible truth-valuations of (24)-(26) consists of all possible truth-valuations with the exception of the one just stated.

Examining truth-valuations of this space, we see that all of them map the schema (27) on *truth*. This reflects the fact that (27) amounts to logically true statements.

$$(\forall x(\boldsymbol{P_1}(x) \rightarrow \boldsymbol{P_2}(x)) \wedge \boldsymbol{P_1}(\boldsymbol{T})) \rightarrow \boldsymbol{P_2}(\boldsymbol{T}) \tag{27}$$

However, for (27) to come out as valid it is essential that we have the space of truth-valuations we do; the valuation mapping (24) and (26) on *truth* and (25) on *falsity* would lead to mapping (27) on *falsity*; and hence should we include this valuation among the admissible ones, we would no longer have (27) as valid.

So we can avoid speaking about mapping parameters on constants, but we cannot avoid dividing truth-valuations between "admissible" and "inadmissible", and this division we in turn cannot make without the recourse to considering possible ways of "dis-abstracting" schemata. The problem of the relationship of schemata to language may be transformed into the problem of delimitation of the space of admissible truth-valuations; but it cannot be dispensed with altogether.

We can fix a space of truth-valuations (the class INT^*) and compare the resulting class of valid statements (or *tautologies*, as valid statements, when arrived at through truth-valuations, are often called) with the class of theorems of an axiomatic system. This is a formal problem and it may be possible to formally prove that the two classes are identical. But to prove this is not to prove material adequacy of the axiomatics. The axiomatics could in such a case be ascertained as adequate only provided that the class of truth-valuations is the class of "naturally admissible" truth-valuations (i.e. that the truth-valuations correspond to material interpretations); and this can never be subject to a formal proof.

Of course, the problem of the specification of the class of admissible truth-valuations is only one side of the coin the other side of which is determining consequence. Valid inferences can alternatively be considered as *limitations of acceptability of truth-valuations*: to say that an inference is valid is to say that a truth-valuation that would verify its antecedent and falsify its consequent is inadmissible. To say that to infer (25) from (24) and (26) is valid means to say that no "possible" interpretation can verify (24) and (26) and falsify (25).

4 Model Theory

4.1 Tarski's Proposal

The proof-theoretic way of capturing consequence was challenged by Alfred Tarski (1936), who in fact denied that the approach would be capable of yielding an adequate account of consequence. Tarski claims that, e.g., to infer *every natural number has the property E* from the infinite number of statements *0 has the property E, 1 has the property E,* ... is legitimate, but in principle not graspable proof-theoretically. He thus urges the necessity of a better account for consequence; the one he proposes is based on the intuitive idea that valid inferences are those which remain valid whatever concrete objects the expressions in them may be considered to refer to.

Tarski starts from the fact that a valid inference remains valid if we replace its extralogical parts by any other grammatically suitable expressions; thus he considers the criterion of consequence stating that a statement X is the consequence of a class K of statements if the following condition holds (ibid., p.7):

> (F) ersetzt man in den Aussagen der Klasse K und in der Aussage X die Konstanten - mit Ausnahme der rein logischen - durch beliebige andere Konstanten (wobei überall gleiche Zeichen durch gleiche ersetzt werden), und bezeichnet man die dadurch aus K erhaltene Aussagenklasse mit "K'" und die aus X erhaltene Aussage mit "X'", so muß die Aussage X' wahr sein, falls nur alle Aussagen der Klasse K' wahr sind.[35]

[35]If, in the sentences of the class K and in the sentence X, the constants - apart from purely logical constants - are replaced by any other constants (like signs being everywhere replaced by like signs), and if we denote the class of sentences thus obtained from K by "K'", and the sentence obtained from X by "X'", then the sentence X' must be true provided only that all sentences of the class K' are true.

This amounts to one of the facts stated in the previous chapter: an instance of logical consequence is invariant to interchange of extralogical constants; in other words, every logical truth is (generally) valid. (However, let us keep in mind that this is, in fact, trivial - validity is invariance w.r.t. interchange of extralogical constants, and extralogical constants are defined as constants that can be interchanged without violating logical truth.)

Tarski proposes considering the inversion of (F); which amounts to what we called *second validity conjecture*[36]. We have seen that the conjectures imply the coincidence of logical truth with general validity, and thus that if accepted as justified, it would allow us not only to reconstruct logical truth as general validity, but to *reduce* the concept of logical truth to the concept of general validity, i.e. to a concept based on truth *simpliciter*.

Tarski is not, however, going to argue for the tenability. He immediately concludes that the inversion of his condition (F) is *not* tenable, that general validity (in our sense) does not necessarily imply logical truth. He mentions no examples, but he remarks that the untenability is due to the possible lack of a sufficient number of extralogical expressions, so it is clear what he has in mind: if the only two constant terms we had were **Ar** and **Pl**, and the only constant predicates **Hu** and **Mo**, then all the substitutional instances of the schema *P*(*T*) would be true, and hence they would be generally valid; they are, however, not logically true[37].

The reason why Tarski acts on all these considerations is that he is convinced that they can show the right way to a criterion of validity. What he insists is that the inversion of (F) *would* be valid if we replaced the notion of substitution by the notion of *satisfaction*. Instead of thinking of parameters as substituted for by constants we should think of them as acquiring varying "referents"; instead of considering which substitutional instances of a schema are true we now should consider which objects satisfy it. Thus, the criterion of validity of *P*(*T*) should no longer be whether it yields a true statement whenever

[36]It is not always clear whether Tarski, when using the term 'logical truth', addresses logical truth in our sense, or necessary truth in a more general sense; hence the inversion of his (F) may actually be closer to the first of our validity conjectures.

[37]Cf. footnote 23. Let us keep in mind that we do not take the constants and their compounds to be elements of a formal calculus, we take them as mere regimentations of their natural language "prototypes". **Hu**(**Ar**), e.g., is the regimentation of *Aristotle is human*; hence it is true, but not logically true.

P and *T* are replaced by suitable kinds of expressions, but rather whether it yields a true statement whenever *P* and *T* are made to refer to suitable kinds of objects.

If we take for granted that every extralogical expression stands for an object, then we can consider this new approach to validity as a generalization of the original, substitutional one. If **Ar** stands for the individual Aristotle, then we can look at the substitution of **Ar** for *T* simply as causing *T* to stand for Aristotle, at the substitution of **Pl** as causing it to stand for Plato, etc. The point of improvement brought about by the shift from substitution to satisfaction is that now it is possible to consider also the reference to objects which happen not to have a name. As there surely exists some property and some individual not having that property, *P(T)* will not come out as valid even if we have no extralogical expressions other than **Hu**, **Mo**, **Pl** and **Ar**.

An assignment of suitable objects to extralogical constants which satisfies a statement (or a group of statements) can be called a *model* of the statement (or of the group). Using this term, Tarski articulates his thesis as follows (ibid., p.9)

> Die Aussage X *folgt logisch* aus den Aussagen der Klasse K dann und nur dann, wenn jedes Modell der Klasse K zugleich ein Modell der Aussage X ist.[38]

We have seen that the extralogical constants of the predicate calculus are terms and predicates. In case of this calculus, satisfaction amounts to an assignment of suitable kinds of objects to terms ("individuals") and to predicates ("properties", "relations") . Tarski considers the relation of satisfaction as something so straightforward that there is no need to clarify it: with respect to such pronouncements as *Johann and Peter satisfy the relation 'X and Y are brothers'* or *2, 3 and 5 satisfy the equation 'x+y=z'* there appears to be no room for any uncertainty. This makes Tarski, and especially many of his followers, believe that satisfaction is a simple uncontroversial relationship to be found within the real world - a belief which, as we are going to argue, is far from plausible.

---

[38]The sentence X *follows logically* from the sentences of the class K if and only if every model of the class K is also a model of the sentence X.

4.2 From Substitution to Satisfaction

Tarski's proposal amounts to an entirely new criterion of logical truth (or consequence), different from the proof-theoretic one, a criterion which can be called *model-theoretic*. Using the conceptual apparatus introduced in the previous chapter we can say that the basic idea of the proposal is to explicate logical truth as general validity (or, in the more general case, necessary truth as validity); but this, Tarski concludes in effect, requires a notion of validity different from the one we have introduced in the previous chapter. The new notion is no longer based on substitution, but rather on reference and satisfaction.

We have called a schema valid if each of its substitutional instances is true; and we have called a statement valid if it is an instance of a valid schema. What Tarski puts forward can be seen as the proposal to the effect of calling a schema valid if it is satisfied by every reference-assignment. Let us speak about *substitutional* validity in case of our definition; and let us speak about *satisfactional* validity in case of Tarski's proposal. Satisfactional validity implies substitutional validity; moreover, if we take reference (instead of substitution) as the basis for defining extralogical constants[39], then we still have logical truth implying general validity. What Tarski insists is that now also the opposite holds: that satisfactional general validity implies logical truth.

Tarski's suggestion may be also viewed as a proposal to the effect that the substitutional notion of instantiation should be replaced by a more general notion, by what can be called *satisfactional* instantiation. Substitutional instances of schemata are statements; and a schema is considered substitutionally valid if each of its substitutional instances is a *true* statement. We can consider introducing a satisfactional notion of instance of a schema - call it a *proto-fact* or a *potential fact* - such that a schema is satisfactionally valid if each of its satisfactional instances is an *(actual) fact*. The substitutional instance of a schema is a statement, a compound of expressions; an instance of $\boldsymbol{P}(\boldsymbol{T})$, e.g., is a compound of two expressions, one being the substitute for $\boldsymbol{P}$,

---

[39]This means that we no longer define a grammatical category as extralogical if its expressions can be freely interchanged without violating logical truth, but rather if the reference of the expressions can be freely varied - within the bounds of a corresponding domain - without violating logical truth.

the other for ***T***. The satisfactional instance, on the other hand, is rather a compound of objects; that of ***P(T)*** is a compound of a property referred to by ***P*** and an individual referred to by ***T***. So ***P(T)*** is substitutionally instantiated by the statement **Hu(Ar)**, whereas it is satisfactionally instantiated by the proto-fact of the individual Aristotle's having the property of being human. Instead of saying that **Hu** and **Ar** substituted into ***P(T)*** yield a true statement we now say that the property of being human and the individual Aristotle *satisfy* ***P(T)***; or that the proto-fact of Aristotle's being human is actual.

Viewed in this way, Tarski's proposal appears to be particularly reasonable: why take the detour of assigning the constant **Ar** to the parameter ***T***, when **Ar** is anyway only a means of presenting the person Aristotle? Why not assign Aristotle directly? Why take the detour of dealing with statements, which express facts, instead of dealing with the facts themselves? What Tarski proposes thus seems to be a clever fix-up making validity into a real tool of account for consequence; and there seems to be no reason not to embrace it.

However, we must realize that Tarski's proposal is an essential point of departure. At the point at which we begin to view the concept of *instance* as a matter more abstract than merely syntactic, at the point when we pass from the concept of substitution to the concepts of reference and satisfaction, we accomplish an essential turn in the character of the entire enterprise of accounting for logical and necessary truth (i.e. of logic); and this turn is almost universally misunderstood. *But if the essential importance and the nature of this turn is not properly recognized, then the possibilities and limitations of a model-theoretic account cannot be understood adequately.*

4.3 The Algebraic Perspective Continued

To make the formal aspects of Tarski's proposal clear, let us return to our algebraic depiction.

Let us recall that we have assumed that we were given a class S (of statements), its subclass T (of true statements), the subclass NT (of necessarily true statements) of T, and the subclass LT (of logically true statements) of NT; and a class Σ (of schemata) plus a mapping INST of Σ into the power set of S. We have defined the subclass V of S by stipulating that $s \epsilon V$ iff there is a $\sigma \epsilon \Sigma$

such that $s\epsilon INST(\sigma)$ and $INST(\sigma)\subseteq T$. Then we assumed that we are also given a subset Σ^* (of general schemata) of Σ and we defined the subclass GV of S by stipulating that $s\epsilon GV$ iff there is a $\sigma\epsilon\Sigma^*$ such that $s\epsilon INST(\sigma)$ and $INST(\sigma)\subseteq T$.

We introduced the class INT amounting to "dis-abstracting" parameters, to mapping them on constants; we assumed that each of the elements of INT maps every schema on a constant statement in such a way that for every $s\epsilon S$ and for every $\sigma\epsilon\Sigma$, $s\epsilon INST(\sigma)$ if and only if $i(\sigma)=s$ for some $i\epsilon INT$. The switch from substitution to satisfaction can now be seen as abandoning considering mappings of parameters on constants in favour of considering mappings of constants on language-external things. This means that instead of INT we now consider a class REF amounting to mapping parameters on some kind of things plus a relation SAT between the elements of REF and schemata; and we assume that the following holds.

for every $i\epsilon INT$ there is an $r\epsilon REF$ such that for every $\sigma\epsilon\Sigma$, r SAT σ
iff $i(\sigma)\epsilon T$ (F21)

Now we can define MOD_{SAT}, $V^{\Sigma}{}_{SAT}$, V_{SAT} and GV_{SAT} - the "satisfactional" analogues of MOD, V^{Σ}, V and GV of the previous chapter.

If $\Sigma'\subseteq\Sigma$ and $r\epsilon REF$, then $r\epsilon MOD_{SAT}(\Sigma')$ iff r SAT σ
for every $\sigma\epsilon\Sigma'$. (D15)
$\sigma\epsilon V^{\Sigma}{}_{SAT}$ iff $MOD_{SAT}(\{\sigma\})=REF$ (D16)
$s\epsilon V_{SAT}$ iff there is a $\sigma\epsilon\Sigma$ such that $s\epsilon INST(\sigma)$ and $\sigma\epsilon V^{\Sigma}{}_{SAT}$ (D17)
$s\epsilon GV_{SAT}$ iff there is a $\sigma\epsilon\Sigma^*$ such that $s\epsilon INST(\sigma)$ and $\sigma\epsilon V^{\Sigma}{}_{SAT}$ (D18)

It is easy to see that (F21) implies (F22) and this further implies (F23)-(F25):

if $MOD_{SAT}(\Sigma')=REF$, then $MOD(\Sigma')=INT$ (F22)
$V^{\Sigma}{}_{SAT} \subseteq V^{\Sigma}$ (F23)
$V_{SAT} \subseteq V$ (F24)
$GV_{SAT} \subseteq GV$ (F25)

Within the substitutional approach we have proved that $NT\subseteq V$ and $LT\subseteq GV$, and we have considered the conjecture that the inclusions are not proper, i.e. that $NT=V$ and $LT=GV$. Tarski rejected this; and his proposal

- if accommodated within our framework - can be considered as amounting to a new kind of conjecturization, which can be articulated as (TC1) and (TC2).

$$V_{SAT}=NT \tag{TC1}$$
$$GV_{SAT}=LT \tag{TC2}$$

However, the satisfactional conjectures (TC1) and (TC2), although seemingly analogous to (VC1') and (VC2'), are essentially different from the substitutional conjectures - whereas (VC1') and (VC2') are real conjectures which can turn out to be true or false (as we assume our ability to specify - empirically - statements that are necessarily true, along with our independent ability to specify - also empirically - statements that are substitutionally valid), (TC1) and (TC2) are simply mere stipulations (because we have - at least so far - no specification of REF and SAT and hence we do not see any independent empirical way to specify the statements that are satisfactionally valid). This means that unless we present a specification of REF and SAT, we must consider (TC1) and (TC2) as mere stipulative constraints put on them.

4.4 Getting Satisfaction Right

This algebraic depiction highlights the essential difference between the substitutional and the satisfactional account: while INST and T, which figure in the substitutional definition of V, are yielded by language itself (via logico-grammatical analysis), REF and SAT, which take their place in the satisfactional version are something that must be determined *in addition to* (and in, fact, independently of) language. Without such an additional determination of REF and SAT the definition of satisfactional validity is essentially incomplete. The basic problem now is *how* to determine REF and SAT.

We have seen that Tarski did not really worry about this problem: for him, satisfaction was a simple naturalistic relation that posed no principal problems. Reference and satisfaction were considered as matters of the way language "hooks on the world"; and they were taken to be straightforwardly accessible to us. This led philosophers like Field (1972) to viewing Tarskian truth theory as part of natural science: as the proposal to the effect of explaining certain obscure naturalistic concepts (namely consequence and necessary truth) by means of other, transparent naturalistic concepts (namely reference and satisfaction) - analogous to, say, the proposal to explain electricity as a stream of certain subatomic particles, namely electrons. According to this view, a theoretician of truth is a natural scientist inquiring into the way words are related to things in the same way as his fellow natural scientist inquires into the processes going on within objects that display electricity.[40]

However, if we try to articulate the formal definition of REF and SAT by way of describing our physical world, we are immediately baffled by severe problems. What kinds of assignments of objects are to constitute REF? What in the real world should count as an object? Obviously individuals and things, but what about events, matters, points of view, or sakes? And what is to count as an individual or as a thing; is a university or a dream a thing? And is Frodo Baggins an individual? And can every kind of object be assigned to every kind of parameter? Moreover, even if we considered such questions concerning REF as somehow settled, it would be hard to see the world as yielding a nontrivial definition of SAT. We can say that a reference-assignment r satisfies *X and Y are brothers* if and only if r(X) and r(Y) are brothers; but this does not seem to be the kind of constraint that could underlie the needed nontrivial formal definition.

So we may try to approach the problem in a different way - if we are unable to give a direct definition, we can attempt at an indirect one; we may try to pool all that we know to hold about reference and satisfaction, treating the resulting pronouncements as constraints which, collectively, characterize REF and SAT implicitly. If we manage to articulate all the constraints which are relevant, then what will result from their combination should really be the desired definition. However, here we face another kind of problem - we can readily see that we can easily formulate a universal satisfactory constraint, which, nevertheless, would bring us no nontrivial gain. The point is that it is clear that if σ is a necessarily valid schema, then σ must be satisfied by every reference-assignment; and hence it is clear that for a concept of satisfactional

[40]Note that such a naturalistic notion of the satisfactional account of consequence leads to grasping logic as "eine Art von Ultra-Physik" as Wittgenstein (1956, §I.8) puts it, "die Beschreibung des 'logischen Baus' der Welt, den wir durch eine Art von Ultra-Erfahrung wahrnehmen (mit dem Verstand etwa)." ["a kind of ultra-physics ... the description of the 'logical structure' of the world, which we perceive through a kind of ultra-experience (with the understanding e.g."]

validity to be able to underlie the adequate account for consequence and necessary truth it is necessary and sufficient to be underlain by a notion of satisfaction that does justice to this. However, if we used this general constraint to articulate the formal definition of REF and SAT; the result would be trivial - we would simply claim that REF is any such set and SAT any such relation between the elements of REF and schemata that for every $\sigma \epsilon \Sigma$

$$\text{r SAT } \sigma \text{ for every r} \epsilon \text{REF iff } \sigma \text{ is necessarily valid} \quad \text{(C)}$$

To find less trivial, more specific constraints, we may consider dividing statements (and consequently corresponding schemata) into groups according to the way they are formed and then consider the problem of reference and satisfaction for each of the groups separately. If we do this, we can see that those statements which are formed by means of the logical operators $\neg$, $\wedge$, $\vee$ and $\rightarrow$ pose no nontrivial problem; their satisfaction conditions are straightforwardly reducible to those of their parts. Thus, a reference-assignment evidently satisfies a schema of the form $S_1 \wedge S_2$ if and only if it satisfies both the schemata S_1 and S_2; and similarly for the other operators. These facts can be turned into straightforward constraints on REF and SAT:

$$\text{r SAT } \neg S \text{ iff not r SAT } S \quad (\text{C}\neg)$$
$$\text{r SAT } S_1 \wedge S_2 \text{ iff r SAT } S_1 \text{ and r SAT } S_2 \quad (\text{C}\wedge)$$
$$\text{r SAT } S_1 \vee S_2 \text{ iff r SAT } S_1 \text{ or r SAT } S_2 \quad (\text{C}\vee)$$
$$\text{r SAT } S_1 \rightarrow S_2 \text{ iff r SAT } S_2 \text{ or not r SAT } S_1 \quad (\text{C}\rightarrow)$$

These rules make it plausible that any definition of REF and SAT working for all statements other than those formed by means of the logical operators can be trivially extended to work also for them. So in view of these constraints, it is enough to give the definition of SAT for the other two kinds of statements: for atomic statements, i.e. statements of the form $P(T_1,\ldots,T_n)$ where P is an n-ary predicate and $T_1,\ldots,T_n$ are terms, and for quantificational statements, i.e. statements of the form $Q\text{x}S^{[\text{x}\rightarrow T]}$ where Q is a quantifier, S a statement, T a term, and $S^{[\text{x}\rightarrow T]}$ is the statement which arises from S by the replacement of T by x. Let us first consider how to define REF and SAT for atomic statements (leaving quantificational statements for the next chapter); and let us first consider purely *logical* truth.

If we restrict ourselves to the investigation of logical truth, then atomic statements pose no problem - no atomic statement is logically true and thus there is no atomic schema that would be logically valid. Thus, it seems that it would be enough if we assumed that for every atomic schema there is both a reference-assignment that satisfies it and one that countersatisfies it; i.e. the only constraint on REF and SAT is the easily satisfiable constraint

$$\text{if } S \text{ is atomic then there are a r,r'} \in \text{REF such that r SAT } S \text{ and not r' SAT } S \qquad (CAt)$$

This means that restricting ourselves to logical truth we can obtain the adequate definition of REF and SAT for atomic statements so straightforwardly, that the situation is trivial. The point is that it is clearly sufficient to have enough of both actual and unactual facts - this alone is enough to guarantee the possibility of defining SAT in such a way that (C*At*) holds. (Note, however, that providing we have enough both of false statements and of true ones, we can do the same via substitutional interpretations; hence in that case we could make do with substitutional validity, we would not need to introduce the satisfactional one. This indicates that as far as atomic statements and logical truth are concerned, the only reason for passing from substitution to satisfaction would be the possible lack of a sufficiently rich repertoire of statements, i.e. our inability to express certain facts.[41])

However, Tarski's project, despite employing the term 'logical truth', seems to be more ambitious than to aim only at an account of what we called *logical* truth. This is, in fact, implied by Tarski's own examples - e.g. the schema *X and Y are brothers* - which are clearly only relevant from the point of view of what we called *necessary* truth. Can the straightforward constraint on REF and SAT, (C*At*), which is enough in case of logical truth, be somehow generalized to the more general case of necessary truth?

The crucial problem raised by the switching from considering purely logical truth to considering necessary truth in general is that we can no longer assume that all atomic statements are independent; we must account for such instances of consequence as (21).

$$\mathbf{Hu}(T) \Rightarrow \mathbf{Mo}(T) \qquad (21)$$

[41] Cf. footnotes 23 and 37.

The indication is that we need a less trivial definition of REF and SAT, which can do justice to instances of consequence beyond instances of logical consequence; thus excluding, e.g., every reference-assignment that would satisfy the schema $\mathbf{Hu}(T)$ and at the same time not satisfy the schema $\mathbf{Mo}(T)$.

For simplicity's sake, let us first consider adding only a single predicate to the class of logical constants[42]. This means that we single out one predicate, say the binary predicate **Br** schematizing the natural language predicate *to be brothers*, and start to consider it as a logical constant whilst maintaining our consideration of all the other predicates as extralogical. In this way we slightly extend the list of instances of consequence to be taken into account; we have to include such instances as $\mathbf{Br}(T_1,T_2) \Rightarrow \mathbf{Br}(T_2,T_1)$. This means that we must exclude some reference-assignments which we previously accepted, e.g. any reference-assignments that would satisfy $\mathbf{Br}(T_1,T_2)$, but not satisfy $\mathbf{Br}(T_2,T_1)$. Thus, we would have to put down several constraints like the following

if r SAT $\mathbf{Br}(T_1,T_2)$, then r SAT $\mathbf{Br}(T_2,T_1)$ (**CBr**)

Shifting more predicates from the list of extralogical constants to that of logical constants necessitates the stipulation of more constraints on REF and SAT. Thus, adding the predicate **Fa** for *being the father of*, we have to add not only the constraints idiosyncratic to **Fa**, such as (**CFa**), but also those capturing the interrelations between **Br** and **Fa**, such as (**CBF**).

if r SAT $\mathbf{Fa}(T_1,T_2)$, then not r SAT $\mathbf{Fa}(T_2,T_1)$ (**CFa**)
if r SAT $\mathbf{Fa}(T_1,T_2)$ and r SAT $\mathbf{Br}(T_2,T_3)$, then r SAT $\mathbf{Fa}(T_1,T_3)$ (**CBF**)

If we take *all* predicates as logical (thereby switching from logical truth to general necessary truth), REF and SAT become far more intricately constrained. We obtain an immensely complex list of new instances of consequence that are to be accounted for; and as a consequence we need an even more immensely complex system of constraints on REF and SAT.

However, there may seem to be a way to turn this menacingly complex list of constraints into a single simple nontrivial constraint; a way that may appear plausible when we take the perspective on model theory usual today. To

[42]At this point, the term *logical constant* might be deemed inappropriate, but we shall not further multiply our termini.

take it, we must do two things: first we must begin to view reference and satisfaction in terms of model structures; and second, we must drop parameters. This is the theme for the following two sections.

4.5 Model Structures

Reference-assignment is considered to be an assignment of some suitable language-external entities to parameters or to extralinguistic expressions; i.e., if we stick to traditional logic, to predicates and terms. If we speak about the objects assigned to terms as about *individuals* and about those assigned to predicates as about *properties* or *relationships*[43] (thereby adopting a mere terminological convention, not making an ontological claim of any kind), then we can say that reference-assignment is a matter of assigning individuals to terms and properties to predicates. Now instead of saying that the schema $\boldsymbol{P}(\boldsymbol{T})$ is satisfied by a reference-assignment r - i.e. by the individual r($\boldsymbol{T}$) and the property r($\boldsymbol{P}$) - we can also say that the individual r($\boldsymbol{T}$) *has* the property r($\boldsymbol{P}$); and instead of saying that $\boldsymbol{P}(\boldsymbol{T}_1,\ldots,\boldsymbol{T}_n)$ is satisfied by r - i.e. by the individuals r($\boldsymbol{T}_1$) through r($\boldsymbol{T}_n$) and the relation r($\boldsymbol{P}$) - we can say that the individuals r($\boldsymbol{T}_1$),...,r($\boldsymbol{T}_n$) *stand in* the relationship r($\boldsymbol{P}$). This way of speaking amounts to "pushing satisfaction inside" the range of reference-assignments: although from the formal point of view we are changing only the way we speak (the new terms are interdefinable with the old ones), we seem to be passing from considering a relation between schemata and reference-assignments to considering internal properties of the domains of entities referred to. Viewed in this way, the domain comes to appear as structured (it appears to be a set of individuals patterned by means of properties and relationships); and model theory begins to appear as relating schemata to certain *model structures*.

Algebraically, the passage from reference-assignments and satisfaction to model structures can be described as follows. Given REF and SAT, for every $r \in$REF we define the model structure $M = \langle D, \langle R^n \rangle_{n \in N} \rangle$ where

[43]From now on we shall use the term *relationship* in this context (i.e. for what is referred to by a predicate of the arity greater than 1); we shall reserve the word *relation* for relations in the mathematical sense (i.e. n-tuples of objects).

(i) D consists of all individuals d such that there is a parametric term T such that $r(T)=d$;

(ii) R^n consists of n-ary relations over D, i.e. of classes of n-tuples of elements of D so that

(ii.i) there is a one-to-one correspondence between the entities (properties or relationships) referred to by n-ary predicates and elements of R^n; i.e. to every entity e such that there is an n-ary parametric predicate P such that $r(P)=e$ there corresponds an element of R^n and vice versa; and

(ii.ii) for every $rl \epsilon R^n$, if $d_1,\ldots,d_n \epsilon D$, then $<d_1,\ldots,d_n> \epsilon rl$ iff r SAT $P(T_1,\ldots,T_n)$, where $T_1,\ldots,T_n$ are some terms for which $d_1=r(T_1),\ldots,d_n=r(T_n)$ and P is a predicate for which rl corresponds to $r(P)$. (We, of course, assume that if r SAT $P(T_1,\ldots,T_n)$ and $r(P)=r(P')$, $r(T_1)=r(T_1')$, ..., $r(T_n)=r(T_n')$, then also r SAT $P'(T_1',\ldots,T_n')$.)

If we now write $\| T \|_M$ for $r(T)$ and $\| P \|_M$ for the relation corresponding to $r(P)$, we can start to say that $P(T_1,\ldots,T_n)$ is *true in M* iff $< \| T_1 \|_M,\ldots, \| T_n \|_M > \epsilon \| P \|_M$; and we can easily prove that a schema is true in M if and only if it is satisfied by r. So every talk about satisfaction by a reference-assignment is formally intertranslatable with the talk about truth in the corresponding model structure. Moreover, if we form the class MSTR of all and only model structures which correspond to the reference-assignments of REF, we can see that a formula is satisfied by all (some) reference-assignments of REF just when it is true w.r.t. all (some) model structures of MSTR; hence also every talk about satisfaction by some or every reference-assignment (hence about satisfiability and about validity) is intertranslatable with the talk about the truth w.r.t. some or every model structure.

If we view the Tarskian apparatus thus in terms of model structures (rather than in terms of reference and satisfaction), we can also consider individual model structures as *possible worlds*; as representations of the various ways our real world may be imagined to be. This seems to pose natural limits to the space of model structures which are "admissible" - only a model structure reflecting what is really possible can be admitted. Thus, it seems to show the straightforward way to define the right boundaries of MSTR - boundaries which are, of course, projectible back to REF and SAT, which were, however, not discernible before we switched the perspective.

4.6 Language as an Algebra

So now we have began to view satisfactional interpretation as comparing schemata with model structures. In this section we show that it is possible to eliminate parameters and to replace schemata by statements, thereby turning substitutional interpretation into what looks like comparing words and things.

To eliminate parameters is - technically - without difficulty: we can simply allow their roles to be played by extralogical expressions. Let us develop this idea in more detail. To this end we shall now reiterate our algebraic depiction of language and of its formalization, which has been very general so far, in slightly more specific terms. A logico-grammatically analyzed language can be seen as a many-sorted algebra $L = \langle \langle C_i \rangle_{i \in I}, \langle R_j \rangle_{j \in J} \rangle$, where each C_i is a grammatical category (i.e. a class of expressions) and each R_j is a grammatical rule (i.e. a function with its domain equal to the Cartesian product of some grammatical categories and its range included in a grammatical category).[44] We can consider the class S of statements as identical with a C_k for certain $k \in I$. Moreover, we assume that the language-algebra is finitely generated by a lexicon; i.e. we assume that there is a collection $W = \langle W_i \rangle_{i \in I}$ of sets, such that $W_i \subseteq C_i$ for every $i \in I$, the union of all W_i's is finite, and W generates the whole algebra L. The elements of W are the constants corresponding to "words" (i.e. further unanalyzed units posited by the logico-grammatical analysis). We assume that expressions of some categories are extralogical, whereas those of the other categories are logical; this means that we assume there is a subset E of I (of indices of extralogical categories).

The introduction of parameters should now be seen as the introduction of a family $P = \langle P_i \rangle_{i \in E}$ of sets of new symbols; we then turn our attention to the algebra L^P generated by $X = \langle X_i \rangle_{i \in I}$, where X_i is P_i if $i \in E$ and X_i is W_i otherwise. L^P is the language arising out of the abstraction of L, the language of parametric expressions and schemata.

The concept of material interpretation, considered in the previous section, can now be analyzed as follows. We may consider an interpretation as a homomorphism from L^P to L induced by a mapping of P into $W^E = \langle W_i \rangle_{i \in E}$; INT is then the class of all such homomorphisms. The switch from substitution

---

[44]For this kind of algebraic view on language see Janssen's (1983) explication of Montague's (1970) notion of universal algebra.

to satisfaction can now be seen as stopping considering mappings of P into W^E and starting instead considering mappings of P into some family $T = \langle T_i \rangle_{i \in E}$ of nonlinguistic objects (of things words "stand for"); instead of the class INT, based on the category-preserving mappings of P into W^E, we now have the class REF of category-preserving mappings of P into T.

The difference now is that there is no straightforward way to extend T to an algebra for the purpose of considering the extension of the mappings of P into T to mappings of L^P into this algebra (in the way we were considering the extension of the mappings of P into W^E to mappings of L^P into L); and that hence in the satisfactional case we cannot proceed in a manner analogous to the substitutional one. Therefore we postulate the relation SAT between reference-assignments and schemata, and in contrast to the substitutional case, where we may define

$$\sigma \epsilon V^{\Sigma} \text{ iff } i(\sigma) \epsilon T \text{ for every } i \epsilon \text{INT}$$

we now define

$$\sigma \epsilon V^{\Sigma} \text{ iff } r \text{ SAT } \sigma \text{ for every } r \epsilon \text{REF}$$

However, let us return to the substitutional case. Let us consider an injective, category-preserving assignment of elements of P to those of W^E; i.e. let us assume that each extralogical constant is associated with a parameter of the corresponding category in such a way that no two different constants are associated with the same parameter. Such association clearly induces an embedding of L into L^P; and assuming that P_i has the same number of elements as W_i for every $i \epsilon E$ (an assumption which does not seem to violate anything), the embedding is an isomorphism of L and L^P. In such a case we can completely ban P and L^P; their roles can be assumed by W^E and L themselves.

This enables us to stop considering homomorphisms from L^P to L and instead consider automorphisms of L - automorphisms induced by category-preserving mappings of W^E into itself. Hence we abandon considering interpretations of parameters by extralogical constants in favour of considering intersubstitutions of extralogical constants. Thus, instead of the class INT we now have the class SUBST of automorphisms of L induced by the category-preserving mappings of W^E into itself.

The very concepts of schema and of instance, which originally caused us to introduce parameters, can now be seen as resting on the concept of substitution: a schema is a statement whose certain parts have been "abstracted away", substituted for by parameters. Let us sketch how it would be possible to build our apparatus avoiding parameters entirely.

The idea underlying the very concept of validity stemmed from our discovery that interchanging certain words does not affect logical truth. Each interchange of extralogical words can be considered as a function mapping statements on statements; so let us assume that the following holds for elements of SUBST

if $s \epsilon LT$, then $f(s) \epsilon LT$ for every $f \epsilon SUBST$ (F26)

Moreover, it seems to be reasonable to assume that SUBST contains the identity mapping and that it is closed under inversion and composition; i.e. we assume

(a) $i \epsilon SUBST$, where $i(s)=s$ for every $s \epsilon S$ (F27)
(b) if $f \epsilon SUBST$, then $f^{-1} \epsilon SUBST$, where $f^{-1}(s)=s'$ iff $f(s')=s$ for every $s \epsilon S$
(c) if $f,g \epsilon SUBST$, then $f \circ g \epsilon SUBST$, where $f \circ g(s)=f(g(s))$ for every $s \epsilon S$

If there is a substitution which maps a statement on another statement, then the two statements may be called *substitutional variants*. Formally we can derive a binary relation SVAR on S as follows:

If $s,s' \epsilon S$, then s SVAR s' iff there is an $f \epsilon SUBST$ such that $f(s)=s'$ (D19)

It follows from (F27) that SVAR is reflexive, symmetric, and transitive; hence

SVAR is an equivalence on S. (F28)

Furthermore, (F26) yields

if $s \epsilon LT$ and s SVAR s', then $s' \epsilon LT$ (F29)

Now we can form the class of equivalence classes of S according to SVAR; and it is just this class that can be identified with Σ^*. If we define schemata in this way, as classes of statements, then the instances of a schema become just its elements; so $\mathrm{INST}(\sigma)=\sigma$.

Then we assume that there is a subclass SUBST' of SUBST yielding the more "fine-grained" analogue SVAR' or SVAR such that we can assume not only (F30) (which follows trivially from (F29)), but also (F31). (In our particular case we can assume that SUBST amounts approximately to interchanging terms as well as predicates, whereas SUBST' terms only.)

$$\text{if } s \epsilon \mathrm{LT} \text{ and } s \text{ SVAR' } s', \text{ then } s' \epsilon \mathrm{LT} \qquad \text{(F30)}$$
$$\text{if } s \epsilon \mathrm{NT} \text{ and } s \text{ SVAR' } s', \text{ then } s' \epsilon \mathrm{NT} \qquad \text{(F31)}$$

We can now consider Σ as resulting from adding the equivalence classes of S according to SVAR' to Σ^*; and it is obvious that Σ, INST and Σ^* defined in this way fulfil the conditions (F1)-(F3) assumed in Section 3.3 ((F1) is trivial, (F2) follows from (F29) and (F3) from (F31)).

This shows that we can take the concept of substitution as the primitive concept on which to base the concept of validity, and that hence we can make do wholly without the concept of parameter. Now similarly we can ban parameters from the satisfactional account of necessary truth; this means that we can stop considering reference-assignments as functions from P to T, and consider them rather as functions from W^E to T. This means that once we are clear about which constants we treat as extralogical and which thus assume the role of parameters in that they become subject of reference-assignment, we can define simply

$$s \epsilon V_{SAT} \text{ iff for every } r \epsilon \mathrm{REF}, \ r \text{ SAT } s \qquad \text{(D20)}$$

Such a turn makes us see reference-assignment not as a matter of "dis-abstraction", but rather as a matter of the study of the relationship between expressions and the things they stand for. Thus model theory comes to be seen as the depiction of the relation of words to things of various possible worlds. This seems to be the essence of Etchemendy's (1990) proposal to consider model theory as *representational semantics*. But this proposal, if considered carefully, can be seen to lead to a blind alley; and it is instructive to see why.

4.7 "Representational Semantics"

In the beginning of his book Etchemendy (1990, p.20) writes:

> When we view a particular theory of relative truth as explicating 'x is true in W', we see it as providing an account of how the *world* wields its influence on the truth values of sentences within a *fixed* language. If characterizing this influence is the aim of our relativized theory of truth, then I will say we are engaged in *representational semantics*.

The idea behind representational semantics seems to be quite simple and appealing: as we know what the world could be like, we know what combinations of objects, properties and relationships could occur and hence which model structures are admissible. What is not so clearly seen if we view the situation from the original Tarskian "interpretational" viewpoint (i.e. in terms of reference-assignments and satisfaction) becomes unmistakably clear if we stick to the "representational" viewpoint (viewing the situation in terms of model structures): as we know that it is not possible that there were an individual which would be human, but not mortal, we cannot admit a model structure containing a human, but immortal individual, and (21) gets trivially sanctioned. Thus, the representational approach seems to be automatically guaranteeing the right space of model structures (the right collections of instances) to accomodate necessary truth. Hence, the approach seems to invite us to replace the vast list of constraints which we previously had to impose on REF and SAT - or on MSTR - by a single, universal constraint: MϵMSTR if it represents a possible state of affairs.

However, is the situation really so straightforward? How do we come to know which model structures are admissible and which not? How do we come to know that one cannot be human and simultaneously not mortal? It is not clear in which sense this question is capable of being answered: we simply *know* it. However, it is evident that to know that a person like Aristotle cannot be human and simultaneously immortal is in fact nothing else than to know that *Aristotle is human* is incompatible with *Aristotle is not mortal* or, in other words, that *Aristotle is human* entails *Aristotle is mortal*. Knowing that we cannot admit a model structure containing an immortal human thus appears to be the same as knowing that (21) holds.

The problem is that although we can indeed delimit the space of admissible model structures by the appeal to what is "in fact" possible and so get it right, if we then use the space to underlie an account for necessary truths, we are susceptible of moving in a vicious circle. Etchemendy writes (ibid., p.77.):

> The class of models is adequate for a representational semantics if it contains a representative for each genuinely possible configuration of the 'empirical' or 'nonlinguistic' world.

Thus, it is our knowledge of what is possible which is constitutive to the space of model structures. However, as knowing what is possible is evidently only one side of the coin the other side of which is knowing what is necessary (*S* is possible iff $\neg S$ is not necessary), and as knowing what is necessary is nothing else than knowing which statements are necessarily true (*S* is necessary iff *'S'* is necessarily true), Etchemendy's claim boils down to a claim that the class of model structures is adequate for a representational semantics if it contains such model structures which adequately account for all necessary truths. We use necessary truth to delimit the space of model structures, and then we use the same space to explain necessary truth.

Representational semantics could be capable of reducing necessary truth to reference and satisfaction only if we were able to know what model structures are possible somehow *independently of knowing what is necessarily true*. However, there is no direct way to enter the space of possible worlds and identify its boundaries save taking the detour via language and necessary truth. To know which worlds are possible is to know what is necessary. To say that there cannot be a world in which there is an individual which is human, but not mortal, is only a *façon de parler* of saying that it is necessary for a human to be mortal, which is in turn the same as saying that *X is human* always entails *X is mortal*. This *façon de parler* may be the most natural one or the most plausible one, it may help us see some aspects of the situation more clearly; however, there is no reason to think that this and only this choice of words would lead to an appropriate kind of theory.

We can put it also in a slightly different way: A model-structure can be considered as a class of instances of individuals having properties and standing in relationships; instances we have proposed to call (proto-)facts. MSTR contains certain classes of such proto-facts and does not contain others (we are, for example, unlikely to allow it to contain a set comprising both the proto-fact of Aristotle's being human and that of Aristotle's being immortal). What representational semantics tells us is that MSTR consists of just those classes of proto-facts which are possible, i.e. which are consistent. However, to know which classes of proto-facts are possible is to know which sets of atomic statements, expressions of proto-facts, are possible. Nevertheless, since statements can be conjoined, knowing which classes of statements are consistent is the same as knowing which single statements are possible; and since statements can also be negated, this is in turn the same as knowing which statements are necessarily true. However, if the *definition* of MSTR rests on necessary truth in this way, MSTR can hardly be used to *explain away* necessary truth.

Representational semantics can be considered as defining MSTR by stipulating that $M \epsilon MSTR$ if and only if it corresponds to a possible state of affairs. However, it turns out that this way of defining MSTR (and consequently REF and SAT) does not differ substantially from the trivial definition (C). Our algebraic articulation therefore leads us to conclude that representational semantics fails to offer a way to define MSTR (REF and SAT) independent of NT. What it does propose is to characterize the space of model structures - and hence reference and satisfaction - simply as *such a class which leads to the class V coinciding with NT.* We may either consider this result to knock down the whole model-theoretic idea, or we may try to look for some alternative value in it. Whichever path we opt for, though, it is the complete model-theoretic account which is at stake, no matter whether we call it "interpretational", "representational", or whatever.

4.8 Model-theory as Criterial Reconstruction

We have seen that there is little hope for a nontrivial reduction of the concept of necessary truth to the concepts of reference and satisfaction, nor to the concept of possible world; there is little hope for defining REF and SAT, or MSTR, independently of NT. Does this mean that the whole idea of model theory is doomed?

In a sense, yes. As long as we considered validity as based on substitution, we could treat it as a well-defined concept which could be taken to reconstruct, and thereby explain away, the concept of necessary truth (if, of course, the two concepts really prove to coincide in extension). This approach is precluded as soon as we pass from substitution to satisfaction - we failed to define satisfactional validity without recourse to necessary truth (or to a concept trivially equivalent to it), and hence we cannot use it to *explain away* necessary truth - on pain of a vicious circle. This implies the failure of the naturalistic conception of model theory, of the conception of "logic as ultra-physics"[45]. However, is this the only way to make sense of model theory?

We argue that it is not. Let us - for a moment - return to the case of proof theory considered in the previous chapter. We have seen that proof theory yields an axiomatic system, the theorems of which are to coincide with necessary truths. Now note that this approach to necessary truth is open to the very same objection which has just appeared to us to undermine model theory. What property of a formal axiomatic system, i.e. of a bundle of statements and inference-rules, makes it admissible as a reconstruction of necessary truth? What makes a statement acceptable as an axiom and another not? Well, it is the capability of the whole of the system to yield the class of theorems "reasonably close" to some chosen class of necessary truths. So if we have revealed model theory to circularly claim *a statement is necessarily true if it is true w.r.t. a space of model structures which does justice to necessary truth*, we now see that proof theory claims, equally circularly, that *a statement is necessarily true if it is provable within an axiomatic system which does justice to necessary truth.* So proof theory seems to be circular in the same way as model theory has just been revealed to be.

However, we have made sense of proof theory despite this apparent circularity - we have seen that its virtue is that it purports to offer a *criterial reconstruction* of necessary truth. Its purpose is to capture the infinity of necessary truths in terms of a finite axiomatic system. Such a system is always subject to assessment with respect to its adequacy, but once we decide to take the adequacy for granted, it gives a criterion; at least till there appears a next substantial challenge to its adequacy. Its point is not in helping us get rid of the concept of necessary truth, but rather only in helping us tame it.

---

[45]See note 35.

And if we can thus make sense of the apparently circular proof theory, we should consider making an analogous sense of model theory; we can take model theory not as a means of explaining away necessary truth, but "merely" as a means of its criterial reconstruction. However, this yields entirely different criteria by which to evaluate the theory. We should no longer be interested in whether model structures - or other means of model theory - faithfully describe something within reality or within our consciousness, we should instead be interested whether (i) the whole of the theory yields a faithful account for necessary truth; and whether (ii) it yields a criterion.

Seen from this vantage point, the circularity of the claim *a statement is necessarily true if it is true w.r.t. a space of model structures which does justice to necessary truth* - as that of *a statement is necessarily true if it is provable within an axiomatic system which does justice to necessary truth* - is only seeming. The point is that the two occurrences of *necessary truth* within the claim mean two different things: while the second means the intuitive necessary truth, i.e. necessary truth as prior to logical reconstruction, the first means necessary truth as it is reconstructed by the theory. The claim should be seen as the conjunction of two subclaims of quite different character: first, there is the purely theoretical and uncontroversial claim *a statement is (called) valid iff it is true w.r.t. a certain space of model structures*, and, second, there is the claim *the space of model structures does justice to necessary truth, i.e. it makes a statement come out valid just in case it is necessarily true* - the assumption of the adequacy of our theory which can never be proved, it simply expresses the belief that the theory is suitable for what it purports to be the theory of.

However, does this not mean that we would approve also such a claim as *a statement is necessarily true if it is necessarily true* as a "theory" of necessary truth? Is this "theory" not directly seen to be adequate? No - although it is adequate, it is not a theory. The point is that it fails to yield a criterion, and it fails to do that so obviously that it cannot even be taken seriously. Our strategy is to reconstruct an intuitively given, and hence "non-criterial", class as a "criterial" one. If we say that a statement is necessarily true simply if and only if it is necessarily true, then we "reconstruct" the original class as the class itself, i.e. again as a "non-criterial" class. If we reconstruct necessary truth as provability within an axiomatic system, we reconstruct it really "criterially" (at least in a sense - because, as we shall see later, we can distinguish between strong and weak criteriality); no matter that the adequacy of the reconstruction will always be subject to assessment and, as the case may be, correction, and that the criterion will - in this sense - depend on the original class. If we reconstruct necessary truth as truth w.r.t. a space of model structures, or satisfaction by a space of reference-assignments, and if we have reasons to believe that the reconstruction is adequate, then the question of the significance of such a reconstruction will be whether the reconstruction introduces a real *criterion*.

The point of criterial reconstruction is making implicit, uncriterial knowledge into an explicit criterion; the criterion being substantiated only insofar as it really captures the pre-theoretical intuition, but being superior to the mere intuitive knowledge in that it is usable as a *norm*. We derive axioms and inference-rules from our notion of necessary truth, and then we use them to *define* what it is to be necessarily true. This is no vicious circle, this is the act of turning a description into a norm.[46] Similarly we delimit the space of admissible model structures, or of admissible reference-assignments, by way of a certain summarization of necessary truths; however, once a criterion is successfully found we then use it to tell us what is necessarily true and what not.

4.9 Formal Soundness and Completeness

However, if model structures are not representations of reality, what are they? The answer to this question should now be clear: model structures are tools internal to a theory of necessary truth, they are "mere" auxiliaries which aid us in constructing a criterion for what it is to be necessarily true. Just as the only indispensable quality of axioms and inference rules of an axiomatic system is that they contribute to the system's criterially reconstructing necessary truth, the only task of model structures is that they add up to a space which yields an adequate criterial reconstruction of necessary truth.

In an article on propositional logic Lukasiewicz and Tarski (1930) described two alternative methods to characterize logical truth: the axiomatic method and what they call the *matrix method*. The principle of the matrix method is that we have a set (in the simplest case the elements of the set are the

---

[46]See Section 2.7.

two truth values) some of the elements of which are in some way distinguished (in the simplest case it is a single element, "the truth") and that we consider recursive assignments of the elements of the set to statements. The recursive assignments are such that every necessary truth is always assigned a distinguished element and that every statement that is always assigned a distinguished element is necessarily true. This is a very instructive way of thinking about model-theoretical tools and it ought to be generalized: model theory for the predicate calculus is only a higher flight of the matrix method - one must not be confused by the fact that in this case we call elements of the matrices "individuals" and "properties".

The model-theoretic approach to characterizing necessary truth is usually considered as something complementary to the proof-theoretic approach. This is, indeed, reasonable: two more or less independent ways of characterization are surely better than a single one. If we have a formal language furnished with both the proof-theoretic and the model-theoretic characterization of necessary truth, we usually call it *sound* if every proof-theoretic truth is also a model-theoretic truth (i.e. if every theorem is valid), and we call it *complete*, if, vice versa, every model-theoretic truth is also a proof-theoretic truth (i.e. if every valid statement is a theorem). In contrast to the concepts of material soundness and completeness introduced in Section 3.5, these concepts of soundness and completeness will be called *formal* soundness and completeness.

If we understood model structures representationally, then we would be forced to see model-theoretic truths directly as necessary truths, and we would be forced to understand formal soundness and completeness as material. From such a viewpoint, a proof of soundness or completeness of a formal calculus, such as Gödel's (1930) proof of completeness of the first-order predicate calculus, would be necessarily understood as amounting to the adequacy of the formal system to reality. However, as we have seen, representationalism is untenable, and formal soundness and completeness cannot be mistaken for material.

In fact, a proof of formal soundness and completeness shows merely that the proof-theoretic and the model-theoretic characterization of necessary truth for a calculus agree. Let us imagine a man going through a forest to a town who is not sure whether he is on the right path. If he meets another man on the same path with the same intention of reaching the town, they may both be surer that their path is the right one, that it will really take them to the town; their agreement is, however, hardly *proof* of the correctness of their decision. Only their arrival in the town could furnish such proof. The model-theoretic and the proof-theoretic explications are like the two men trying to reach the town (intuitive notion of necessary truth): that they meet corroborates the adequacy of both of them, it is, however, not a *proof* of either.

And, obviously, the coincidence of the original class of intuitive necessary truths with its reconstruction (whether proof-theoretic, model-theoretic, or whatever) can *never* be subject to anything like a formal proof[47]. We can compare two classes by comparing either their elements, or their alleged criteria; however, if the classes are infinite, only the second possibility remains. This second possibility then clearly applies only to classes which are "criterial", which are defined via such or another kind of criterion; and this is not the case with the class of intuitive necessary truths[48].

Tarski seemed to believe that his generalization of instantiation, his replacement of substitution by reference and satisfaction, meant only the replacement of one kind of material interpretation by another kind. But this is not true: this move inevitably means the replacement of the material by the formal interpretation (and hence of material by formal soundness and completeness). It is thus the shift from external properties of the system to internal ones[49].

---

[47]It is sometimes claimed that adequacy of a formal system can be proven formally along the lines of Kreisel (1967). But if what Kreisel presented can be understood as a formal proof, then it must be considered to amount to conicidence of two *formal* notions; and if it is to be understood to amount to material adequacy, it immediately loses the unquestionable correctness of a formal proof. To think otherwise means, as Brouwer (1907, p.76) puts it, to make "the mistake which so many people made, thinking that they could reason logically [i.e. formally, in our terms - J.P.] about other subjects than mathematical structures built by themselves."

[48]If some neural machinery could be found in our brain which would be responsible for "producing" necessary truths or consequences, then it could possibly in some way be compared with the proof-theoretic or the model-theoretic one; but this is a mere science-fiction.

[49]In the sense of Carnap (1950a).

4.10 Model Theory vs. Proof Theory

A criterion for entities of an infinite range is a rule which yields (specifications of) all and only entities of the range. A criterion can be called *strong* if it, moreover, allows us to definitely decide whether a given entity does or does not fall into the range. It is evident that if we have a weak criterion for belonging to a given range plus a weak criterion for not belonging to that range, we can combine them into the strong criterion for belonging to the range. If we have a weak criterion of "consequencehood" plus a weak criterion of "nonconsequencehood", we have a strong criterion of "consequencehood": if we are to decide whether an inference is valid, we start to parallelly generate the list of instances of consequence and that of instances of nonconsequence, and sooner or later we must encounter the instance in question in one of the lists.[50]

A theory which purports to deliver a criterial reconstruction of a range of entities or a property amounting to such a range, must be checked for two qualities. First, it must be checked whether the theory really yields a criterion, because a theory may deliver a reconstruction which fails to be really criterial. An extreme example of a theory of consequence which would fail to yield a criterion would be the "theory" made up by the statement *an n-tuple of statements and a statement constitute an instance of consequence if the latter is the consequence of the former*; or the trivial "theory" of necessary truth mentioned in Section 4.7. Such theories reconstruct the uncriterial class which is to be accounted for as again an uncriterial class, i.e. as a class which is in the same need of being accounted for as the original one; such theories do not explicate, but rather only paraphrase. Second, if a theory does yield a criterion, it is useful to check whether the criterion is weak or strong; theories which yield strong criteria are in general superior to those which yield only weak ones.

---

[50]One of the possible formal reconstructions of criteriality available inside set theory is *recursivity*: what we call weak criteriality could be understood as recursive enumerability, and what we call strong criteriality as recursivity. However, let us keep in mind that we are not in a formal theory, let alone formal set theory.

The difference between proof theory and model theory can now be seen in that whereas proof theory offers a criterion of consequencehood (which is not necessarily strong); model-theory is after an account for both consequencehood and nonconsequencehood. Thus, to the extent to which it were possible to consider model theory to yield real criteria, it would be superior to proof theory in that its criterion of consequencehood would be strong. The problem is, however, that it is far from clear that this is really possible.

Model theory gives prescriptions how to decide whether a given reference-assignment satisfies a schema in case we know whether the valuation satisfies its parts. It tells us, for instance, that

$$\text{r SAT } S_1 \wedge S_2 \text{ iff r SAT } S_1 \text{ and r SAT } S_2 \qquad (\text{C}\wedge)$$

or, in terms of model structures

$$S_1 \wedge S_2 \text{ is true w.r.t. M } (\| S_1 \wedge S_2 \|_{\text{M}} = 1) \text{ iff } S_1 \text{ is true w.r.t. M } (\| S_1 \|_{\text{M}} = 1) \text{ and } S_2 \text{ is true w.r.t. M } (\| S_2 \|_{\text{M}} = 1) \qquad (\text{C}\wedge\text{'})$$

or it tells us that

$$\text{r SAT } \mathbf{Mo}(T) \text{ if r SAT } \mathbf{Hu}(T) \qquad \textbf{(CHM)}$$

in terms of model structures

$$\mathbf{Mo}(T) \text{ is true w.r.t. M } (\| T \|_{\text{M}} \in \| \mathbf{Mo} \|_{\text{M}}) \text{ if } \mathbf{Hu}(T) \text{ is true w.r.t. M } (\| T \|_{\text{M}} \in \| \mathbf{Hu} \|_{\text{M}}) \qquad \textbf{(CHM')}$$

Some of such prescriptions can be directly turned into rules of characterization of the class of instances of consequence (or necessary truth), into rules of proof. They then give rise to what is usually called rules of natural deduction[51]. Thus, to posit (C$\wedge$) or (C$\wedge$') is clearly the same as to posit the three inference rules ($\wedge_1$)-($\wedge_3$).

$$S_1 \wedge S_2 \Rightarrow S_1 \qquad (\wedge_1)$$
$$S_1 \wedge S_2 \Rightarrow S_2 \qquad (\wedge_2)$$

[51]See Beth (1955).

$$S_1, S_2 \Rightarrow S_1 \wedge S_2 \qquad (\wedge_3)$$

Similarly, to posit (**CHM**) or (**CHM'**) is the same as to posit (21).

If model-theory were restricted to this kind of rules, as it is in case, e.g., of the classical propositional calculus, or in the case of predicate calculus without quantifiers, then doing model theory could be considered as simply a peculiar way of doing proof theory.

The main problem now is whether we can find similar constraints on REF and SAT to do justice to quantificational statements. If we managed to find them, then we could conclude that model theory is sound; but also that it is not an enterprise really substantially different from proof theory.

However, as we are going to argue in the following chapter, this is, in fact, not the case; model-theoretic capturing of quantificational statements does not yield constraints that could be directly turned into rules of proof. This may either lead us to conclude that model theory in general simply does not succeed in delivering a criterion, and hence that it does not lead to the real characterization of consequence which logic seeks; or we may conclude that what model theory delivers *is* a kind of criterion, although not a criterion in a straightforward sense. (As we are not inside a formal system, we cannot base the decision of whether something is a criterion on formal rules - this would require to articulate a criterion of criteriality and consequently a vicious circle).

5 Quantification

5.1 Quantificational Statements

We have seen that from the viewpoint of model theory, statements formed by means of logical operators can be simply eliminated with the help of a couple of recursive constraints reducing satisfaction of such statements to satisfaction of their substatements. The idea flowing from this is to handle quantificational statements analogously; to find analogous recursive constraints for quantificational statements. Attempting this, though, we immediately face a problem - quantificational statements do not in general contain substatements, at least not in the overt way those formed by means of logical operators do.

However, a quantificational statement like $\forall x\mathbf{Hu}(x)$, can be considered as resulting from a statement like $\mathbf{Hu}(\mathbf{Ar})$ by generalizing on its subject; and so there seems to be a sense in which the latter can be considered as contained in the former. Thus, we may consider looking at $\mathbf{Hu}(\mathbf{Ar})$ as a part of $\forall x\mathbf{Hu}(x)$. In this way we come to view every quantificational statement as containing substatements; and we may consider the relation between the truth value of the former and those of the latter. The most straightforward attempt to articulate the relation yields the following constraints

$$\text{r SAT } \forall x S^{[x\to T]} \text{ iff r SAT } S^{[T'\to T]} \text{ for every term } T' \qquad (\mathrm{C}\forall)$$
$$\text{r SAT } \exists x S^{[x\to T]} \text{ iff r SAT } S^{[T'\to T]} \text{ for some term } T' \qquad (\mathrm{C}\exists)$$

If we consider the constraint (C$\wedge$), introduced in the previous chapter, as amounting to the reduction of the truth value of the statement $S_1 \wedge S_2$ to the truth values of the statements S_1 and S_2, then we can consider (C$\forall$) as amounting to the reduction of the truth value of the statement $\forall x S^{[x\to T]}$ to the truth values of the set of statements containing $S^{[T'\to T]}$ for every term T'. In both cases we want to reduce satisfaction of a more complex statement to satisfaction of multiple simpler statements.

However, unlike the straightforward constraints of the kind of (C$\wedge$), the constraints (C$\forall$) and (C$\exists$) are problematical; they pose problems of two kinds.

First, the class of statements to which $\forall x S^{[x \to T]}$ is to be reduced may well be infinite, and hence not treatable so straightforwardly as a finite class. Second, and this seems to be worse, the constraint appears to be sure to work in one direction only - if r SAT $\forall x S^{[x \to T]}$, then it appears to be sure that r SAT $S^{[T' \to T]}$ for every term T'. The other direction, however, is problematic: it seems that if there are individuals which are the value of r for no term, then we can have r SAT $S^{[T' \to T]}$ for every term T', but not r SAT $\forall x S^{[x \to T]}$. Let us call the first problem *the problem of infinity (of instances)* and the second *the problem of lack (of instances)*.

The insufficiency of the substitutional account for quantification as manifested by *the problem of lack* is parallel to that of the substitutional account for necessary truth, and also the standard way to handle it is analogous - it is the shift from substitution to satisfaction. The usual way is to replace the substitutional constraints by their satisfactional analogues of the following kind (where the non-bound variable x is understood as a parameter and $r^{[x \Rightarrow d]}$ is the reference-assignment that is just like r save that $r^{[x \Rightarrow d]}(x)=d$):

$$\text{r SAT } \forall x S^{[x \to T]} \text{ iff } r^{[x \Rightarrow d]} \text{ SAT } S^{[x \to T]} \text{ for every individual d} \qquad (C\forall')$$
$$\text{r SAT } \exists x S^{[x \to T]} \text{ iff } r^{[x \Rightarrow d]} \text{ SAT } S^{[x \to T]} \text{ for some individual d} \qquad (C\exists')$$

5.2 Quantification and Necessary Truth

The problem of the model-theoretic account for quantificational statements is, we have seen, one chapter of the problem of the account for the whole of language (other chapters being - if we keep ourselves restricted to traditional logic - the account for the statements formed by means of logical operators, and the account for atomic statements). However, now we have also seen that the problems posed by the account for this particular kind of statements are closely parallel to those posed by model theory in general. No wonder - the basic pronouncements of model theory are quantificational. This is to say that although out of the statements which are the subject of model theory, only some are quantificational, the core of the theory itself is constituted by claims the nature of which is quantificational.

We saw that model theory was motivated by the idea of explaining necessary truth as general truth, of reducing necessary truth to truth *simpliciter* of a number of instances. We proposed to consider a statement necessarily true if and only if it is an instance of a valid schema, and to call a schema valid iff all its instances are true. Wishing to avoid the dependence of necessary truth on the resources of a particular language, we then switched from substitution to satisfaction and thus from statements to proto-facts; the final proposal was to consider a statement necessarily true iff the corresponding schema is satisfied by every reference-assignment, i.e. if all the proto-facts to which the schema might be considered to refer are actual.

In the case of general quantificational statements the situation is evidently similar. The original proposal (C$\forall$) amounted to considering a universally quantified statement true iff all statements on which it is considered to generalize are true; in the modified variant (C$\forall$') it is proposed to consider the statement true if all proto-facts which it is considered to summarize are actual.

Let us, for the sake of illustration, compare the approach to necessary truth with that to quantificational truth by considering the statements (28) and (29).

$$\mathbf{Hu(Ar)} \rightarrow \mathbf{Mo(Ar)} \tag{28}$$
$$\forall x(\mathbf{Hu}(x) \rightarrow \mathbf{Mo}(x)) \tag{29}$$

Let us first look at (28). The "provisional", substitutional definition of validity led to reducing the necessary truth of (28) to the truth of all substitutional instances of the schema $\mathbf{Hu}(\mathit{T}) \rightarrow \mathbf{Mo}(\mathit{T})$, i.e. to the truth of $\mathbf{Hu(Ar)} \rightarrow \mathbf{Mo(Ar)}$, $\mathbf{Hu(Pl)} \rightarrow \mathbf{Mo(Pl)}$, etc. The "improved", satisfactional definition of validity now leads to reducing the necessary truth of (28) to the actuality of all corresponding satisfactional instances, i.e. to the actuality of Aristotle's being human implying Aristotle's being mortal, of Plato's being human implying Plato's being mortal, etc. What all of this amounts to is the claim that a statement s has a property (namely the property of being necessarily true) just in the case that all relevant instances (originally the substitutional, later the satisfactional instances of the relevant schema) have another, "simpler" property (originally that of being true, later that of being actual).

It is readily seen that the satisfaction constraints put on quantificational statements are of the very same kind. Let us consider (29). (C$\forall$) states that a reference-assignment r satisfies (29) if and only if r satisfies $\mathbf{Hu(Ar)} \rightarrow \mathbf{Mo(Ar)}$, $\mathbf{Hu(Pl)} \rightarrow \mathbf{Mo(Pl)}$, etc. However, as (29) contains no parameters, r satisfies (29) if and only if (29) is true; and the same holds for **Hu(Ar)→Mo(Ar)**, **Hu(Pl)→Mo(Pl)**, etc. Thus, (C∀) applied to (29) can be seen as amounting to the reduction of the truth of (29) to the truth of **Hu(Ar)→Mo(Ar)**, **Hu(Pl)→Mo(Pl)**, etc. Similarly for (C∀') - applied to (29) it can clearly be looked at as claiming the reduction of the truth of (29) to the actuality of the proto-facts of Aristotle's being human implying Aristotle's being mortal, of Plato's being human implying Plato's being mortal, etc. Hence satisfactional constraints on general quantificational statements can be seen as amounting to the claim that a certain statement s has a property (namely the property of being true) just in the case that all of a certain range of instances (originally substitutional, later satisfactional), which are "simpler" than s, have another property (originally that of being true, later that of being actual).

The shared pattern thus is the following

s is $\mathcal{P}$ iff all instances of a certain range associated with s are $\mathcal{P}'$

or, schematically,

$$s \in \mathcal{P} \text{ iff for every } s' \in R_s,\ s' \in \mathcal{P}' \tag{30}$$

The point of the pattern is to reduce a "stronger" claim to multiple "weaker" claims - in the substitutional case we reduce either a "stronger mode" of truth (namely necessary truth) to a "weaker mode" of truth (namely truth *simpliciter*), or the truth of a more complicated statement to the truth of its "parts", i.e. of "simpler" statements.

If we drop talk about individual instances, and instead talk about whole substitutional interpretations and reference-assignments, we can say that (28) is necessarily true if the schema **Hu(*T*)→Mo(*T*)** is verified by every interpretation, or satisfied by every reference-assignment; and similarly for (29). This amounts to claiming that an s is $\mathcal{P}$ if and only if s stands in certain relation R to every element *e* of a collection C (of interpretations, or of reference-assignments); hence the schematic pattern now is (30').

$$s \in \mathcal{P} \text{ iff for every } e \in C,\ s\ R\ e \tag{30'}$$

We may further underline the parallelism between quantificational truth and validity-based reconstruction of necessary truth by drawing out the quantificational character of the latter. Thus we can introduce a schematic expression for claims about validity and instead of writing *(Hu(T)→Mo(T)) is satisfied by everything T may refer to* we can schematically write, e.g.,

$$\Pi T(\mathbf{Hu}(T)\rightarrow\mathbf{Mo}(T)) \tag{31}$$

However, we must remember that this is a *new* way to employ schematic means. Here we are not using schematic means to regiment pre-theoretical statements nor to articulate types of such statements, but we are using them to shorten pronouncements of the theory - therefore we have used specifically new symbols, symbols not previously employed. This means that we must keep in mind the difference between (31), which regiments the theoretical claim *every T satisfies Hu(T)→Mo(T),* and (29), which regiments the pre-theoretical, "natural" statement *every human is mortal.* (31) is defined to be true if and only if $\mathbf{Hu}(T)\rightarrow\mathbf{Mo}(T)$ is valid, and hence, if we accept validity as an adequate reconstruction of necessary truth, it is true if and only if (28) (and any other statement of the form $\mathbf{Hu}(T)\rightarrow\mathbf{Mo}(T)$) is necessarily true.

Thus, the necessary truth of (28) is the same as the truth of (31); and hence to account for the necessary truth of (28) is to account for the truth of the quantificational statement (31). From this angle, the switch from substitutional to satisfactional validity, i.e. from V to V_{SAT}, displays itself as the same move as the switch from substitutional to satisfactional quantification, i.e. the switch from (C∀) to (C∀')[52]. The common problem to be overcome by the switch is that we worry whether there are sufficiently many substitutional instances; that, in terms of (30), even if every one of the substitutional instances associated with s does have $\mathcal{P}$', s may lack $\mathcal{P}$. The solution is simply to add instances. Whenever there is a case of s not having $\mathcal{P}$, despite all the instances associated with s having $\mathcal{P}$', we understand it as the indication that there "in fact" is an unnoticed relevant instance, an instance which does not have $\mathcal{P}$'. The switch from substitution to satisfaction is the method to secure the needed supply of such instances.

[52]Kripke (1976) has drawn the attention to the basic distinction between substitutional quantification which can be construed as referential quantification over expression, and quantification which is *irreducibly* substitutional. This distinction does not play any substantial role here - the difference between what we call substitutional and satisfactional quantification can be well construed as the difference between referential quantification over expressions and that over objects.

However, the price which must be payed for the unlimited supply is that instances become creatures of darkness: they are no longer clearcut and independent of that which they are invoked to account for. We often think about quantificational truth in terms of examples such as "Everybody in this room wears glasses." In this case, the verification of the statement is the matter of inspecting several persons. However, the situation is quite different in the general case: in the cases of *every human is mortal* or *everything green is extended* there can be no similar going through the relevant instances.

5.3 "Dis-necessitation"

Besides mirroring the core problem of the whole model theory, the problem of quantificational statements seems to bring in something new. The concept of validity, as introduced by model theory, can be seen as the vehicle for reducing the necessary truth of a given statement to the simple truth of a set of other statements, or to a satisfaction by a set of reference-assignments. Now the truth of such a set of statements, or the satisfaction by such a set of reference-assignments, appears to be reducible to the truth of a single quantificational statement; hence we might consider reducing necessary truth to the simple truth of a single - quantificational - statement. We saw that it is the same for (28) to be necessarily true as for (31) to be true *simpliciter* - for (28) is reconstructed as necessarily true iff $\mathbf{Hu}(T)\rightarrow\mathbf{Mo}(T)$ is satisfied by every reference-assignment, and $\mathbf{Hu}(T)\rightarrow\mathbf{Mo}(T)$ is in turn satisfied (by definition) by every reference-assignment iff (31) is true *simpliciter*; hence the necessary truth of (28) may be considered as reducible to the simple truth of (31). This can lead us to articulate a principle facilitating the turning of necessary truth into simple truth, a principle of "dis-necessitation":

> A universal quantificational statement - like (31) - is true if and only if an instance of its matrix - like (28) - is necessarily true.

However, if we formulate the principle in this way, then we must realize that it is trivial. We must not forget the difference between the character of statements like (31) and of those like (29). The fact that the statement *every human is mortal* and its regimentation (29) is true is a pre-theoretical fact; it is a datum which is to be accounted for by the theory. The same holds for the fact that (28) is necessarily true. (31), on the other hand, is not a symbolization of a pre-theoretical statement - it is a pronouncement yielded by the theory, a *theoretical* statement.

This means that to say that we *reduce* the validity (28) to the truth of (31) could be misleading: it seems to mean that we have discovered a dependency between two pre-theoretical facts. However, saying that (28) is valid just when (31) is true is not making a substantial reduction, and it is in fact not making any substantial claim whatsoever; it is rather merely repeating the definition of the truth value of (31). This means that saying that (31) is true is not an explication of the fact that (28) is valid, it is merely its re-statement. "Reducing" the validity, and hence the necessary truth, of (28) to the truth of (31) is equally trivial as "reducing" its validity to the truth of the statement ***Hu(Ar)***$\rightarrowtail$***Mo(Ar)*** *is valid.*

However, this is, of course, not the only way to render the idea of reducing necessary truth to truth *simpliciter* of a quantificational statement. We may consider equating the necessary truth of a pre-theoretical statement like (28) not to the truth of a *theoretical* quantificational statement like (31), but rather to that of a *pre-theoretical* quantificational statement, like (29). This kind of reduction is, indeed, nontrivial and substantial - it claims the coincidence of two of the pre-theoretical facts which we are to account for.

So we may consider an alternative formulation of the "dis-necessitation" principle

> A universal quantificational statement - like (29) - is true if and only if a statement on which it "generalizes" - like (28) - is necessarily true.

We have seen that the coincidence of the validity of (28) with the truth of (31) is trivial, that it is simply a trivial consequence of the way we decided to use the sign $\mathbf{\Pi}$; but if we understand the "dis-necessitation" principle rather as claiming that the validity of (28) coincides with the truth of (29), then we make a nontrivial claim. Understood in this way, the principle claims that (28) is necessarily true if and only if (29) is true, or, de-schematized, that *If Aristotle is human, then he is mortal* is necessarily true if and only if *Every human is mortal* is true; and it in fact implies that there is no substantial difference between (31) and (29), and more generally between $\mathbf{\Pi}$ and $\mathbf{\forall}$.

Provided this principle is warranted, we really have a non-trivial "dis-necessitation": the study of necessary truth of at least some statements can be replaced by the study of truth *simpliciter* of corresponding quantificational statements; and we could - at least to a certain extent - think about turning the pursuit of necessary truth (i.e. logic) into the pursuit of quantificational truth. However, we must realize at least two important points related to this claim - one minor and the other major.

The minor point is that not for every necessary truth (and thus not for every theoretical quantificational truth) is there a plausible corresponding natural quantificational truth. Thus, although we may consider (28) as identifiable with (29), there appear to be no really natural counterparts of such theoretical statements as, e.g.,

$$\Pi T \Pi P_1 \Pi P_2 ((P_1(T) \wedge \forall x(P_1(x) \rightarrow P_2(x))) \rightarrow P_2(T));$$

and more generally no really natural counterparts of theoretical statements with quantifiers binding predicates.[53] This is not to say that these theoretical statements cannot be de-schematized and articulated informally (we can, indeed say something like *if an individual has a property and everything that has that property has another property, then the individual has the other property*); it is to say that there would hardly be any such natural statements if there were no logical theory.

The major point then is that when speaking about validity, we distinguished between two ways of considering truth and consequently two concepts of validity. The first way amounts to the actual state of our world, and accordingly to be valid means to have all instances actually true. The second one is more relaxed and amounts to saying that to be valid means something like to have all instances imaginable as true. We argued that it is only the second notion that is applicable within logic, and it was this notion of validity that we embraced.

---

[53]It seems to be precisely this fact, namely that in natural language we can easily generalize on terms, but not on predicates, which can be considered as one side of the coin whose other side is the fact that we are inclined to think about the referents of terms in a much more "thingish" way than about those of predicates, i.e. that we take individuals as more straightforward inhabitants of the real world than properties.

Now the same ambiguity carries over to quantificational statements. *Every human is mortal* can be understood in two ways: first, as claiming that every *actually living* human being is mortal (which may be the result of pure chance, of the contingent fact that, e.g., no immortal human has been born yet); or, second, that it is, so to say, in the nature of humans to be mortal[54], that the concept of being human includes the concept of being mortal. On the first reading, the statement could be possibly refuted by observation - it does not exclude the possibility of encountering an immortal human. On the second reading it *does* exclude this possibility - it means that if we encounter an immortal being, we shall definitely not take it to be human.

It is only the second way of understanding quantificational statements that enables the accommodation of necessary truth to quantificational truth. There is a danger of accepting the reduction principle on the basis of the sound intuition that quantificational statements understood in the second way imply the necessary truth of the statements on which they generalize, and than applying the principle to quantificational statements understood in the first way.

Etchemendy (1990, p.98) articulates what we called the "dis-necessitation" principle under the name of *reduction principle* and he concludes (ibid., p.107):

> When we equate the *logical* truth of a sentence with the *ordinary* truth of a universal generalization of which it is an instance, we risk an account whose output is influenced by facts of an entirely 'extralogical' sort. Clearly, the question of whether the sentence *If Leslie was president then Leslie was a man* is a logical truth does not depend on the sorts of historical facts that determine the truth or falsity of the generalization ∀*x[if x was president then x was a man]*.

What is the relation between *If Leslie was president then Leslie was a man* and *Every president was a man*? Well, it depends on the way we understand the latter statement. If we understood it as stating that it is somehow in the nature of a president to be a man, then its truth would imply the necessary truth of *If Leslie was president then Leslie was a man*; however, if

[54]That a human is mortal κατ᾽ εἶδος, "according to its nature", as Aristotle would put it.

we take it to say simply that there has been no woman president so far, then its truth is evidently not satisfactory to guarantee the necessary truth of the statements on which it can be seen as generalizing. We may distinguish the two readings of *Every president was a man* also by considering them as related to two different domains of quantification: to the domain of the individuals of the real world, and to the domain of all possible, imaginable individuals.[55]

Etchemendy evidently takes the domain of reference of the quantified statement for granted: he takes it to be the real world. This is in accordance with Tarski's declared intention; on the other hand, in combination with the reduction principle it leads to such absurd consequences that we can hardly believe that this was really what Tarski had in mind. We can hardly believe that he would not notice such a serene inconsistency in his thought as to imply that necessary truth and consequence depend on the momentary existence of individuals with some properties in the world (e.g. female presidents), and that it is thus something which may change at any moment.

If we consider *∀x[if x was president then x was a man]* as a contingent statement, it will be false (due to the existence of women presidents such as Corazón Aquino or Violetta Chamorro). But even if it were true, by chance (e.g., if we understood it, as Etchemendy does, as contextually restricted to *American* presidents), it would surely not guarantee the validity and hence necessary truth of *If Leslie was president then Leslie was a man*. On the other hand, if we consider *∀x[if x was president then x was a man]* to express the subsumption of the property of being a man under the property of being a president, then we could argue that the statement cannot be true because *If Leslie was president then Leslie was a man* is not true necessarily - for there is an at least "imaginable" or "hypothetical" president who is not a man.

It may be tempting to view the coincidence of the truth of a theoretical quantificational statement with the necessary truth of a pre-theoretical statement, or with the truth of a pre-theoretical quantificational statement, in the "representational" way - to view (31) as explicating the necessary truth of (28) or the truth of (29) by means of summarizing the world's wielding its influence on statements as "depicted by" individual reference-assignments. However, we have concluded in Chapter 4 that such a naturalistic view of reference is hardly tenable. It makes no nontrivial sense to say that (28) is necessarily true or (29)

[55]Let us note that it is precisely this distinction which Carnap (1955) considers indicative of, and establishing, the distinction between extension and intension.

is true *because* all individuals satisfy **Hu(*T*)→Mo(*T*)**, it is more adequate to say instead that the necessary truth of (28) or the truth of (29) implies that there exist only such individuals which satisfy **Hu(*T*)→Mo(*T*)**.

To be able to understand (31) as stating something over and above the necessary truth of (28) or the truth of (29), we would need to be able to find out about its truth independently of finding about the necessary truth of (28). We determine the truth value of (31) by means of inspecting the space of admissible reference-assignments, but as we saw in Chapter 4, the delimitation of the space is one side of the coin whose other side is necessary truth.

5.4 Tarski's Argument

Let us now turn our attention to *the problem of infinity*, the problem that the constraints enabling us to reduce satisfaction of quantificational statements to those of their "parts" invoke infinitistic means. The nature of this problem can be most clearly seen from the example presented by Tarski himself in his crucial paper.

Tarski argued that the proof-theoretic approach is unable to account for the fact that the statements 'N has the property E' for every natural N together entail the statement 'All natural numbers have the property E'. Let us reproduce the relevant part of Tarski's paper (1936, pp.2-3) literally:

> Schon vor mehreren Jahren habe ich ein Beispiel, übrigens ein ganz elementares, einer derartigen Theorie gegeben, die folgende Eigentümlichkeit aufweist: unter den Lehrsätzen dieser Theorie kommen solche Sätze vor wie:
> A_0. *0 besitzt die gegebene Eigenschaft E,*
> A_1. *1 besitzt die gegebene Eigenschaft E,*
> u.s.w., im allgemeinen alle speziellen Sätze von der Form:
> A_n. *n besitzt die gegebene Eigenschaft E,*
> wobei 'n' ein beliebiges Symbol vertritt, das eine natürliche Zahl in einem bestimmten (z. B. dekadischen) Zahlensystem bezeichnet; dagegen läßt sich der allgemeine Satz:
> *A. Jede natürliche Zahl besitzt die gegebene Eigenschaft E,*

> auf Grund der betrachteten Theorie mit Hilfe der normalen Schlußregeln nicht beweisen. Diese Tatsache spricht, wie mir scheint, für sich selbst: sie zeigt, daß der formalisierte Folgerungsbegriff, so wie er von den mathematischen Logikern allgemien verwendet wurde, sich mit dem üblichen keineswegs deckt. Inhaltlich scheint es doch sicher zu sein, daß der allgemeine Satz *A* aus der Gesammtheit aller speziellen Sätze A_0, A_1, ..., A_n, ... im üblichen Sinne folgt: falls nur alle diese Sätze wahr sind, so muß auch der Satz *A* wahr sein.[56]

Tarski thus claims that it is evidently valid to infer the statement *All natural numbers have the property E* from the infinite set {*0 has the property E*, *1 has the property E*, ...} of statements, and that the proof-theoretic approach is incapable of accounting for this.

Let us analyze Tarski's argument in detail. What Tarski claims is that the statement *All natural numbers have the property E* is true whenever all statements of the set {*0 has the property E*, *1 has the property E*, ...} are true. However, what does it mean to say that all statements of such an infinite set are true? To say that a definite statement is true is a kind of paraphrastic way of asserting the statement itself. To claim that the statement *1 has the property E* is true is to say that 1 has the property E. However, to say that the statement *N has the property E* is true for every N is no paraphrase: it is simply impossible to assert the infinite number of statements directly.

Let us suppose we are to put down an instance of consequence. If the instance amounts to the fact that something follows from the truth of the statement *1 has the property E* (i.e. from the fact that 1 has the property E),

---

[56]Some years ago I gave a quite elementary example of a theory which shows the following peculiarity: among its theorems there occur such sentences as: A_0. *0 possesses the given property P*, A_1. *1 possesses the given property P*, and, in general, all particular sentences of the form A_n. *n possesses the given property P*, where 'n' represents any symbol which denotes a natural number in a given (e.g. decimal) number system. On the other hand the universal sentence: *A*. *Every natural number possesses the given property P*, cannot be proved on the basis of the theory in question by means of the normal rules of inference. This fact seems to me to speak for itself. It shows that the formalized concept of consequence as it is generally used by mathematical logicians, by no means coincides with the common concept. Yet intuitively it seems certain that the universal sentence A follows in the usual sense from the totality of particular sentences A_0, A_1, ..., A_n, ... Provided all these sentences are true, the sentence A must also be true.

then we may consider it to follow either from the statement *'1 has the property E' is true* or simply from the statement *1 has the property E*. However, if we are to state that something follows from the fact that the statement *N has the property E* is true for every N, then there is no alternative: we have no choice but to write *'N has the property E' is true for every N.*

We cannot list all the statements of an infinite set, so the articulation of the truth of all its members inevitably boils down to the truth of such or another "meta"statement. What Tarski himself in fact says is that it is impossible to account for the fact that the "meta"statement *'N has the property E' for every natural N* entails the statement *All natural numbers have the property E.* And the employment of the "meta"statement is no shortcut: there is no way of avoiding it[57].

However, what is the real difference between *'N has the property E' for every natural N* and *All natural numbers have the property E*? To state *'P(x)' is true for every x* is nothing else than to say *P(x) for every x* and this is what gets regimented to $\forall x\ P(x)$; the only substantial difference is that (from the viewpoint of the predicate calculus) the first statement is a *meta*statement, while the last one is a statement proper. What makes the difference between *P(x) for every x* and $\forall x\ P(x)$ is that moving from the former to the latter we cross the boundary between the metalanguage and the object language. It is the crossing of this boundary which does the job of explication; eschewing this boundary would mean to turn the explication into a kind of tautology. However, it seems that to eschew it is precisely what we must do if we are to stop restricting ourselves to formal languages and if we want to apply the Tarskian approach to natural language.

---

[57]The whole problem is in fact only a special case of the general problem of the infinite. It has become usual to consider the difference between a finite and an infinite set on a par with the difference between two finite sets with different numbers of elements. This is, however, quite misguided. As Wittgenstein (1953, §208) puts it: "Es ist zu unterscheiden: das 'usw.', das eine Abkürzung der Schreibweise ist, von demjenigen, welches dies *nicht* ist. Das 'usw. ad inf.' ist *keine* Abkürzung der Schreibweise." ["We should distinguish between the 'and so on' which is, and the 'and so on' which is not, an abbreviated notation. 'And so on ad inf.' is *not* such an abbreviation."] More about this in Chapter 9.

5.5 Statements vs. Metastatements

So, in order to be able to understand the nature of the Tarskian approach to quantification properly, we must next clarify the significance of the distinction between the object language and metalanguage, which plays such an important role in his project. For a formal language, the situation is clear - the object language is the language under study, and the metalanguage is the natural, or another formal, language we use to speak about it. But what if we take, as we do, *natural* language, or some kind of its symbolic rendering, as the object language? Restricting oneself to a fixed part of natural language and leaving the rest to be used as the metalanguage means treating natural language as formal - the naturalness of natural language consists precisely in the fact that it is in no way gerrymandered. If what we are after are philosophical questions about the nature of truth or consequence, or about the possibilities and limitations of logical analysis, this approach can be misguiding. These problems concern pre-formal facts and the ways these facts get formalized; and to deal with them on the basis of turning natural language into another formal language means begging the question. We have to take the whole of language as the object language; and there is no room left for a real metalanguage.

Nevertheless, there does seem to be a way of making sense of the opposition between the linguistic and the metalinguistic without posing an undesirable restriction; we might, it seems, consider it as the distinction between the theoretical and the pre-theoretical, as we have introduced it above. We have stressed the important distinction between *natural*, *pre-theoretical* statements and *theoretical* statements - the former are subject to the theoretical account, the latter are brought to life only by the theory, they are its vehicles. However, this distinction is not the same as the Tarskian one; and it is important to beware of understanding the concept of theoretical statements in an unwarranted way.

Let us return to our previous examples:

$$\mathbf{Hu(Ar)\to Mo(Ar)} \quad (28)$$

$$\forall \mathbf{x(Hu(x)\to Mo(x))} \quad (29)$$

$$\Pi T(\mathbf{Hu}(T)\to\mathbf{Mo}(T)) \quad (31)$$

Two possible ways of viewing theoretical statements, such as (31), were indicated in the previous section. First, there is the most straightforward possibility, namely to understand them as summarizations of pre-theoretical facts, of truths and necessary truths. We may take (31) to express the claim that all statements of a certain range, namely all statements of the form **Hu**(*T*)→**Mo**(*T*), are necessarily true. However, let us stress that (31), if understood in this sense, is a mere summary expression of the necessary truth of (28) and its substitutional variants, it is not an explanation of *why* they are necessarily true. If we take the necessary truth of (28) to coincide with the simple truth of (29), then (31) can be taken to amount also to the truth of (29), but again, it does not tell us *why* (29) is true.

Second, we may accommodate some theoretical statements to certain pre-theoretical statements; we may understand (31) as in fact saying the same as (29). This, though, involves the trivializing the theoretical character of (31); and in general it would render - at least some - theoretical statements redundant.

Neither of the two possibilities allows us to understand a claim to the effect of the equivalence of a pre-theoretical quantificational statement with its theoretical companion, i.e. the equivalence of, e.g., (29) with (31), as allowing us to explain away the truth of the former by means of the latter. The first reading of theoretical quantificational statements renders such claim as stating the equivalence of the pre-theoretical quantificational statement with the necessary truth of the statements on which it generalizes, i.e. the truth of (29) with the necessary truth of (28) and the like. The second reading renders the claim completely vacuous; under this reading the claim simply spells out the way we understand theoretical quantificational statements like (31).

However, there may seem to be a *third* possibility, namely to understand (31) as somehow directly depicting how the world is, and thereby explaining *why* (28) is necessarily true and/or (29) true. The actual state of the world, as expressed by (31), is taken to make (28) necessarily true and/or (29) true. However, this way of viewing theoretical statements assumes that the aim of logical theory is to provide a description of the factual world better than the description of the same world provided by the pre-theoretical language; i.e. that logic is in fact an "ultra-physics". What is assumed is in fact that something can be explained by merely replacing **∀** by **Π** - for the passage from (29) to (31) consists in precisely this alone. From this point of view, the sense of constraints of the kind of (C**∀**') turns out to be questionable. They purport to explicate the symbol **∀** by means of the words *for every* (or their regimentation, such as **Π**).

This would be fine if we understood $\forall$ as a component of a formal language, but it becomes senseless once we consider $\forall$ as a mere regimentation of *for every* - in such a case it simply spells out the tautological claim that what holds for every item holds for every item.

The difference between metalanguage and theoretical statements as we have introduced them lies thus, we can say, in that the former purports to be a more direct presentation of the world than the language under study, whereas the latter are merely summarizations of facts about this language.

To summarize: the Tarskian approach to quantification amounts to the step from a quantificational statement to the set of simpler statements to which it is to be reduced. However, this step is unavoidably the step merely to the *description* of such a set - for (i) the elements of the set are infinite in number (hence *the problem of infinity*) and (ii) they are not all necessarily available (*the problem of lack*). This means that we cannot do more than to move from one quantificational statement to another quantificational statement, making the step appear nontrivial by calling the latter statement a *meta*statement.[58] Thus, although Tarski's results concerning truth and consequence are indisputable if considered in the context of formal languages (they reveal many important and interesting facts about such languages - e.g. the impossibility of defining truth in a formal language by means of the very language), they are to a large extent worthless when applied to natural language. Moreover, they may be even harmful - their problem is that they are very strong in invoking the illusion that they have explained something.

5.6 Quantification and "Logical Atomism"

Logical atomism, the view of the world propagated during the early part of this century especially by Russell (1914) and Wittgenstein (1922), assumes that there is a basic supply of elementary, mutually independent (proto-)facts and that the current state of the world can be seen as the matter of which of these elementary facts are actual and which not; and that there is a corresponding basic supply of elementary statements expressing these facts, so

[58]See also Peregrin (forthcoming d).

that every state of the world can be described by way of distributing truth values among these elementary statements. Each state of the world is thus uniquely determined by a distribution of truth values among elementary statements; and, due to the independence of elementary facts, also, vice versa, every such distribution uniquely determines a possible state of the world. This view implies that the truth value of every statement is uniquely determined by the truth values of the elementary statements.

The way in which logical positivists combined this view with earlier versions of empiricism and made it basic to their methodology of science then suggested the assumption of the principal coincidence of the order of language not only with that of the world, but also with that of our knowledge. This led to perceiving the existence of the basis of elementary statements from which all other statements can be deduced as tantamount not only to the existence of the basis of elementary facts from which all other facts can be constituted, but also to the existence of the basis of elementary pieces of our knowledge ("observations") from which all other knowledge can be inferred.

We can consider the triune structure of our world/language/knowledge as in Figure 1. If this diagram is considered as amounting to the world, the nodes must be assumed to stand for facts and the lines to express a relation of constitution; if we take it to amount to language, we must take the nodes to stand for true statements and the lines to express consequence; and, finally, if we think of it as amounting to our knowledge, we must assume the nodes to amount to elementary pieces of knowledge and the lines to a derivation or inference. The important feature of the diagram is the bottom basis of the elementary facts/statements/knowledge-pieces - it is the fundament that firmly supports everything else; and everything else is in fact only a way of its "unfolding". Every point of the whole edifice is somehow implicitly contained in the points of the basis; in the case of language, the truth value of every statement is fixed by the truth values of the elementary statements.

If we take the viewpoint of language, then it is precisely this aspect of the picture that proves to be tricky. It is assumed that the truth value (or satisfaction conditions) of every non-elementary statement is reducible to the truth values (or satisfaction conditions) of some simpler statements, and that hence everything is reducible - by way of recursion - to the basis. This assumption works well in case of statements built by means of logical operators, but it does not work in the case of quantificational statements.

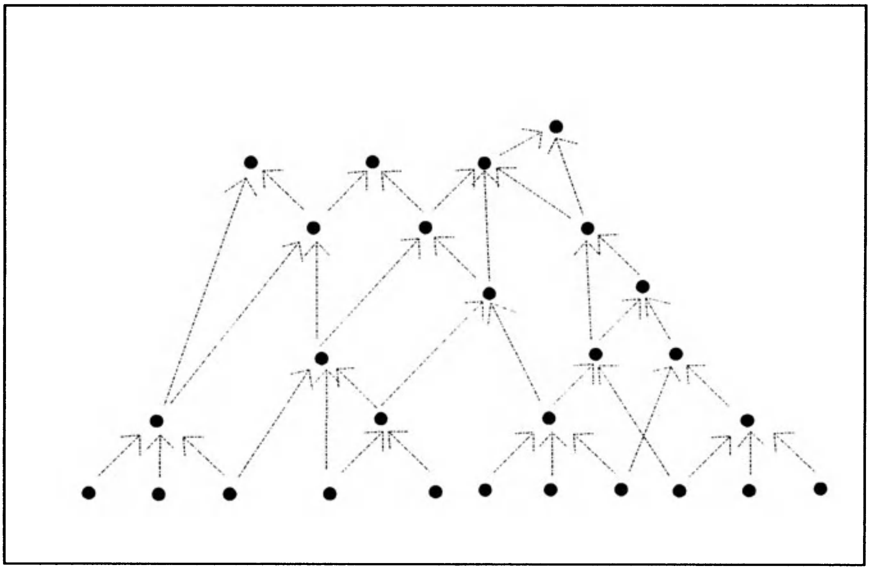

Figure 1

One way to handle this, the way adopted by Wittgenstein (1922), is to accommodate quantifiers to logical operators; to treat general quantification as disguised conjunction, and existential quantification as disguised disjunction. This means to consider the statement (29) as a mere shortcut for the conjunction of all the available statements of the shape **Hu**(T)→**Mo**(T). However, such a solution brings about the two problems we discussed above, *the problem of infinity* and *the problem of lack*. The former consists, we saw, in the fact that if we accommodate general quantification to conjunction (and existential to disjunction), then we do not make do with finite conjunctions (and disjunctions). We are forced to disavow any substantial distinction between finite and infinite sets of statements, which, as we have seen in the previous section, changes the character of the whole enterprise. Indeed the realization of the untenability of this disavowal formed one of the principal reasons leading to Wittgenstein's repudiation of logical atomism.[59]

Moreover, even if we put up with going beyond finite means, there is no guarantee that there will always be all the statements needed to accomplish the

[59]See Kenny (1972, Chapter 6).

reduction; that whenever there will be a false universal statement, there will also be the statement spelling out a counter-instance. It seems to be possible in principle that the statement **Hu(*T*)→Mo(*T*)** were true for every available term *T*, but the statement (29), nevertheless, false. The doctrine of logical atomism offers a solution: if there is an instance lacking within language, then there is surely one within the world - whereas the language-edifice can be imperfect, the world-edifice cannot. This is, of course, true, but equally of course, it brings in nothing nontrivial. The falsity of ∀x**P**(x) indeed means that there is an individual which is not **P**, but to say that there is an individual which is not **P** is to state ∃x¬**P**(x) which is the same as to state ¬∀x**P**(x). Without the crutch of the language/metalanguage boundary it yields the trivial fact that falsity implies truth of negation.

However, in one respect the shift from the falsity of ∀x**P**(x) to the truth ∃x¬**P**(x) may be seen as involving something substantial - it leads us to the vantage point from which we view quantificational truth in terms of reference and existence. Let us explain this in detail.

5.7 Truth and Reference

Let us consider the statements (A) through (D) and the truth values these statements can acquire.

Hu(Ar) (A)
Hu(Pl) (B)
Hu(Pl) ∧ Hu(Ar) (C)
∀x**Hu**(x) (D)

All the distributions of truth values among these four statements which can be imagined (disregarding all dependencies among them) are listed in the following table:

	(A)	(B)	(C)	(D)
1	T	T	T	T
2	T	T	T	F
3	T	T	F	T
4	T	T	F	F
5	T	F	T	T
6	T	F	T	F
7	T	F	F	T
8	T	F	F	F
9	F	T	T	T
10	F	T	T	F
11	F	T	F	T
12	F	T	F	F
13	F	F	T	T
14	F	F	T	F
15	F	F	F	T
16	F	F	F	F

Table 6

The following obvious instances of consequence can be stated for (A)-(D):

Hu(**Ar**), **Hu**(**Pl**) $\Rightarrow$ **Hu**(**Pl**) $\wedge$ **Hu**(**Ar**)
Hu(**Pl**) $\wedge$ **Hu**(**Ar**) $\Rightarrow$ **Hu**(**Ar**)
Hu(**Pl**) $\wedge$ **Hu**(**Ar**) $\Rightarrow$ **Hu**(**Pl**)
$\forall$x**Hu**(x) $\Rightarrow$ **Hu**(**Ar**)
$\forall$x**Hu**(x) $\Rightarrow$ **Hu**(**Pl**)

To state the first of these inferences is to say that **Hu**(**Pl**) $\wedge$ **Hu**(**Ar**) is true whenever both **Hu**(**Ar**) and **Hu**(**Pl**) are, hence it is to render the third and the fourth row of the table of possible truth-valuations as inadmissible, to strike them out of the table. To state the next inferences means to strike out the rows

9, 10, 13, 14; 5, 6, 13, 14; 9, 11, 13, 15; and 5, 7, 13, 15. Hence we obtain the reduced table of the admissible truth-valuations as follows:

	(A)	(B)	(C)	(D)
1	T	T	T	T
2	T	T	T	F
3	T	F	F	F
4	F	T	F	F
5	F	F	F	F

Table 7

Now let us assume that (A)-(D) constitute the whole of language. According to the doctrine of logical atomism, the truth values of all the statements should be uniquely determined by those of the elementary statements, namely of (A) and (B). This is the case of (C); however, it is not the case of (D): the lines 1 and 2 of Table 7 show that the truth value of (D) can be independent of those of (A) and (B).

If we insist on the doctrine of logical atomism, we are forced to conclude that the set of elementary statements we have is not the "right" one (if it were, then it would determine the truth values of the whole of language). We may consider taking (D) as another elementary statement; however, (D) is not independent of (A) and (B), its truth forces the truth of both (A) and (B); whereas elementary statements should be mutually independent. So the only solution seems to be to assume that there is some further elementary statement, different from (A) and (B), which is somehow "inovert" or "latent", that there is a "latent" term, say **X**, and hence a "latent" statement (E) which "explains" the variability of the truth value of (D) which is independent of (A) and (B).

Hu(X) (E)

The truth value of (E) can now be set up in such a way that the table of admissible truth-valuations of (A), (B), (D) and (E) appears as follows (we omit

(C) as it is no longer interesting for us); the truth value of (D) is now uniquely determined by those of (A), (B) and (E).

	(A)	(B)	(E)	(D)
1	T	T	T	T
2	T	T	F	F
3	T	F	T	F
4	T	F	F	F
5	F	T	T	F
6	F	T	F	F
7	F	F	T	F
8	F	F	F	F

Table 8

If we strike out the third column and leave out repeating lines, the table reduces to the previous one. Thus, the space of acceptable truth-valuations of (A), (B) and (E) is left unchanged, but the admissible truth-valuations are now in the one-to-one correspondence with possible distributions of truth values among elementary statements.

What disturbs the happy-end is the dubiosity of the *ad hoc* introduction of **X**. In fact, we saved the situation by tampering with the language we originally wanted to describe. What is the status of **Hu(X)** which we have so unscrupulously added? An answer may be that it is something that has been latently present with the original language and that we have now visualized. That the existence of the **X** such that **Hu(X)** is false is implicit to the fact that ∀x**Hu**(x) is false while both **Hu(Ar)** and **Hu(Pl)** are true; and that to add the statement (E) is merely to make this explicit.

However, we may want the abstract manoeuvre to be explained in somewhat less abstract words. To this end we can tell the story of "going from language to the world". The step from Table 7 to Table 8, we might say, consists in ceasing to consider the truth values of statements and instead to consider the existence of facts. The T's and F's within the cells of Table 8 should not be read as truth values of statements standing in the heading of the corresponding columns, but rather as the existence of facts depicted by the headings. Thus, the introduction of (E) does not amount to extending language, but rather to pointing out a fact whose depiction is beyond the reach of language.

However, if we view the situation in this way, we must not forget that it is a metaphor. For we did not introduce (E) as the result of an investigation of the world - but rather as the result of our effort to cope with a problem posed by Table 7 in view of the doctrine of logical atomism. X is not someone we encountered in the world and then recorded into Table 8; we have posited him to guard against the failure of logical atomism - he is a poor fall guy who is sure not to be human unless everybody is human. This is, however, not to say that the introduction of X has nothing to do with the world, that it is merely the product of our deliberation. The introduction was forced by the truth values of the statements of the considered language, and these truth values were in turn forced by the world. The model structure is not an iconic picture of the world, it is the redrawing of the map of the world as constituted by the truth values of statements.

We can say that what this example lets us see is how our talk gives birth to a "world" (of the talk); however, we should beware of taking such a metaphor literally. What is in fact at stake can be understood simply in terms of reorganization of information, of expressing the content of a theory in a different, somehow more illuminating, way. Let us consider the theory T consisting of the following statements

$$\exists x(\mathbf{P}(x) \wedge \mathbf{Q}(x))$$
$$\neg\exists x(\mathbf{P}(x) \wedge \neg\mathbf{Q}(x))$$
$$\exists x(\mathbf{R}(x) \wedge \neg\mathbf{Q}(x))$$
$$\forall x\exists y(\mathbf{Q}(x) \rightarrow (\neg\mathbf{P}(y) \wedge \mathbf{Q}(y)))$$
$$\neg\exists x(\mathbf{P}(x) \wedge \mathbf{R}(x))$$
$$\forall x(\mathbf{P}(x) \vee \mathbf{R}(x))$$

This theory is quite incomprehensible; and it becomes much more comprehensible when we realize that its "minimal model" consists of three individuals, say A, B and C, such that A has the properties **P** and **Q**, but not **R**; B has the properties **Q** and **R**, but not **P**; and C has the property **R**, but not **P** and **Q**. This is to say that every model structure with respect to which T is true must contain such three individuals, and the model structure which contains nothing more than them verifies T. We can express this by means of the following table

	P	Q	R
A	+	+	-
B	-	+	+
C	-	-	+

This table conveys us "the existential part" of the information conveyed by T; however, it articulates the information more aptly than T itself and it makes its comprehension much easier[60]. However, this must not divert us from seeing the fact that showing that the table expresses the "minimal model" of T is tantamount to proving that T entails the following theory T', which is - in a certain sense, which could be, of course, made more precise - prominent from the viewpoint of T's "existential import":

$$\exists x(\mathbf{P}(x) \wedge \mathbf{Q}(x) \wedge \neg \mathbf{R}(x))$$
$$\exists x(\neg \mathbf{P}(x) \wedge \mathbf{Q}(x) \wedge \mathbf{R}(x))$$
$$\exists x(\neg \mathbf{P}(x) \wedge \neg \mathbf{Q}(x) \wedge \mathbf{R}(x))$$

This indicates that what is crucial about the step from the original theory to the properties-distribution table is not that it would mean "switching from the expressions to what they refer to" - the step can well be understood as a step from one theory to another, in a certain sense more comprehensible, theory.

[60]For an interesting discussion of the ways in which such a rearrangement or representation of information can influence grasping and understanding see Dennett (1991, Chapter 10).

5.8 Reference and Existence

In the previous chapter we have concluded that it is untenable to consider necessary truth as reducible to validity defined independently; we have stated that validity is only a criterial restatement of necessary truth. Thus, it is not adequate to consider the space of admissible reference-assignments as determining necessary truth; necessary truth is rather, the other way around, constitutive to the space.

The same now holds if the model-theoretic approach is applied to the truth of quantificational statements. We may say that if $\forall x\mathbf{Hu}(x)$ is true, then it is because there is no reference-assignment which would not satisfy $\mathbf{Hu}(T)$, i.e. that there is no individual which would not be human. But again, as in the case of necessary truth, to say that there are no inhuman individuals is *nothing more* than to say that $\forall x\mathbf{Hu}(x)$ is true. A model-theoretic account of the truth of $\forall x\mathbf{Hu}(x)$ is not its explanation, it is merely its restatement.

Thus, not: *Everybody is human* is true *because* there are only humans; but rather: that *Everybody is human* is true says that there are only humans[61]. This means that to display *what there is* is not to explain *what is true*; displaying *what there is* is rather only a way of paraphrasing or reconstructing *what is true* in a new way. From this vantage point, *to be* means *to figure within an adequate "model-theoretical" account for truth*; or, in Quine's more aphoristic terms, *to be is to be a value of a variable*. Quine's (1963, pp.13-14) famous verdict is that "...a theory is committed to those and only those entities to which the bound variables of the theory must be capable of referring in order that the affirmations made in the theory be true."

What is interesting about the model-theoretical account for necessary truth is that it brings about a really new point of view, a point of view from which we begin to look at necessary truth and at the patterns of consequence in terms of a universe of discourse. As with any nontrivial shift of viewpoint, this approach can reveal to us something which we could not see from other perspectives; but it does not lead us really "beyond language".

If we know *what there is*, then we know how to determine *what is true*. This is in fact what Tarskian model theory specified, at least for the predicate

[61]Cf. Wittgenstein (1922, §3.1432).

calculus and its generalizations. But what about the inverse direction, which is, as we have concluded, primary? Could we explicitly state *what there is* given *what is true*? Is there a way to construct an explicit universe of discourse, say in the framework of set theory, given the class of true statements?

The domain to which we can be said to be committed is the "simplest" domain with respect to which all the true statements of our talk (especially all the quantificational statements) come out as verified. We saw an example of how to go from a theory to what it is "model-theoretically committed to" in the previous section. A general account of the way to construct the corresponding domain out of the theory is possible; it was envisaged by Henkin (1949; 1950) in his proofs of the completeness theorems. The basic idea is to assume an individual for each constant term and for each true existential statement, and to assume the identity of every two individuals whose identification does not tamper with the truth values of the original statements.

To identify the universe a theory is "about", to discover what the terms of the theory "refer to", is, however, not to pick up some objects existing in the world independently of language[62]. To say that one speaks *about something* is to say that there is no way of making a "model-theoretic sense" of his claims without assuming a universe containing this something. An object referred to by an expression is not something independent of language, it is a reification of the contribution the expression brings to truth. As Quine puts it elsewhere[63]:

> I see reference, reification, and ontology no longer as a goal of science, but rather as a spin-off of quantification and the variables, these being in turn a mere technical aid in forging logical links between observation sentences and theoretical sentences.

Or, as articulated more explicitly by Stekeler-Weithofer (1986a, pp.173-174):

---

[62]Notice that to understand Henkin's algorithm of model-construction it is not essential to consider it as a matter linking the words to what they stand for; it can be understood in terms of extending languages and theories. See Peregrin (forthcoming d).

[63]Barrett and Gibson (1990, p.115).

Die Rede von den Gegenständen der Welt ist nicht sprachunabhängig zu verstehen; unsere üblichen und uns vertrauten Redeweisen *konstituieren* vielmehr sowohl die Rede von den 'Bedeutungen', als auch von der 'Referenz' der Wörter, sowohl die *grammatischen* Gegenstände, als auch das, was wir 'Dinge' der (physischen oder sogenannten 'realen') Welt ('Wirklichkeit') nennen.[64]

5.9 The Non-reducibility of Quantification

For the usual way of logical formalization of language, as exposed in the forerunning chapters, the concept of quantifier is crucial. Whereas logical connectives, predicates and terms can be considered as arising from a more or less direct schematization of natural language, the correspondence of the two quantifiers $\exists$ and $\forall$ to the resources of natural language which they purport to capture is much less straightforward[65]. We saw that it is the introduction of

[64]Our talk about objects in the world cannot be understood independent of language. Rather our well known and common form of speaking *constitute* the meaning of our talk about meaning on the one hand, the reference of the names on the other, the internal objects of abstract talk ("grammatical entities") as well as what we call "things" in the physical or real world or "reality".

[65]The relation is usually also obscured by the confusion about the role of variables (and by the confusing variables and parameters). We have seen in Chapter 2 that from the purely formal viewpoint there may be good reasons to consider variables as fully-fledged elements of the logical schematism; however, this should not prevent us from seeing that from the substantial viewpoint they are mere unnecessary auxiliaries. We have seen that variables of the first-order predicate calculus could be seen as a way of codifying grammatical rules. Thus variables were used to distinguish two different ways in which $\exists$, $\forall$ and **Ad** can be put together, the ways usually being expressed by means of variables as (12'a) and (12'b). The same could be accomplished by means of other technical tools; thus instead of (12'a) and (12'b) we could equivalently use (12'c) and (12'd); or (12'e) and (12'f) (cf. Section 2.4):

$$\forall x \exists y \mathbf{Ad}(x,y) \tag{12'a}$$

$$\exists y \forall x \mathbf{Ad}(x,y) \tag{12'b}$$

$$\mathbf{Ad}(\forall^1,\exists^2) \tag{12'c}$$

$$\mathbf{Ad}(\forall^2,\exists^1) \tag{12'd}$$

$$\forall\ \exists\ \mathbf{Ad}(\ ,\ ) \tag{12'e}$$

these two quantifiers where the classical logical formalizations becomes really nontrivial.

We also saw that it is quantification which marks the point of departure between the ways of proof theory and model theory. As long as we limit ourselves to the propositional calculus, or the predicate calculus without quantifiers, then model theory and proof theory can be seen as simply two variants of the same enterprise. It is only when we start to consider quantification that they begin to appear as two really separate approaches to logic.

The truth of a quantificational statement is usually considered to be reducible to the truth of multiple simpler statements or to satisfaction by multiple reference-assignments - in the same way as necessary truth is usually considered to be reducible to truth *simpliciter* of multiple statements or to satisfaction by multiple reference-assignments. The idea, we saw, was that of the reduction of a "stronger" truth to multiple "weaker" truths. As long as the "weaker" truths to which the "stronger" truth is being reduced are explicit and finite in number, the idea is straightforward and indisputable. We have seen the problems that occur if we abandon the firm ground of language and if we pass from substitution to satisfaction. However, even if we kept to the substitutional approach, the whole idea would become problematic as soon as we accept that the truths to which quantification is reduced are infinite.

We have seen that for the substantiation of the whole model-theoretic enterprise the distinction between object language and metalanguage is crucial. But such a distinction is a purely technical matter - our natural language is certainly not really split in such a way. And if we do not take such a distinction for granted, then what Tarski urges is the possibility of inferring a (quantificational) statement (*Every natural number has the property E*) from another quantificational statement (*'N has the property E' holds for every N*). The point is that the quantification within the former statement is considered as a quantification of the object language, whereas that in the latter statement as

---

$$\forall\ \exists\ \mathrm{Ad}(\ ,\ ) \qquad (12'f)$$

This is a *general* way in which variables can be viewed. (For the exposition of the way to do predicate calculus without variables see Quine, 1960b; a similar approach can be taken to variables in any "reasonable" calculus). This indicates that we could consider variables simply as auxiliaries helping us to codify complicated rules.

a quantification of the metalanguage. What is in question is thus in fact not a step from multiple "weaker" truths to a single "stronger" truth, but rather the shift of quantification from metalanguage to object language[66]. And such a shift loses its sense if what we are addressing is not a formal language, but rather a natural language or its logical schematization. The point is that (C$\forall$') and (C$\exists$') amount to explaining the expressions $\forall$ and $\exists$ in terms of the expressions *every* and *some*; and if we take $\forall$ to be the mere regimentation of *every* and $\exists$ of *some*, then they are hopelessly circular. (C$\forall$') and (C$\exists$'), unlike the constraints like (C$\wedge$), cannot be turned into rules of proof; but it is problematic to infer from this fact that these constraints supply something which rules of proof cannot provide (and that, more generally, model theory provides for characterization of consequence over and above that which is provided by proof theory). The point is that the job done by them is in fact that of robbing Peter to pay Paul.

To say that something is a natural language is to say that it is the "ultimate metalanguage", that it cannot be folded into another language; as soon as we step out of the language to address it by another language, we are not addressing the language itself but only its fossil. To address the phenomenon of truth, as Tarski did, on the basis of the assumption that for every particular language which we might consider we always have a metalanguage is like addressing the phenomenon of death assuming that any particular death we might address is always merely in a movie which real people watch. The proponents of the doctrine of logical atomism tried to secure their stance by considering the world itself as a kind of final, foolproof metalanguage; but their attempt was hopeless - there is no *theoretical* way to go from language to the world, the only possible way, the *practical* one, is not available for a theoretician.

We feel that quantification in its least trivial sense amounts to infinity. And indeed, a quantifier is, as Goldfarb (1979) put it, "that constituent of formal systems which allows us to speak about the infinite". A quantificational statement transcends its finite nature towards infinity; it is, however, a serious misunderstanding to believe that the infinite can be seized in some essentially more explicit way. Quantification is an "indirect" way of speaking about the infinite; however, there is no "direct" way. It is like a thing and its perception: the perception is not the thing itself, but if we accept that there are such entities

---

[66]Cf. Stekeler-Weithofer (1986a, p.18).

as perceptions, then we must conclude that it is always inevitably the perception, and not the thing itself, that we "have"; and it would be foolish to conceive of a more direct contact with the thing itself amounting to "having the thing" in the same sense in which we have the perception.

6 Truth

6.1 Tarski's Notion

In the previous two chapters we discussed the possibilities and limitations of approaches to capturing necessary truth. Are there some analogous ways of capturing truth *simpliciter*? After all, necessary truth seems to be a subspecies, a special case of truth - is not every account for necessary truth thus a special case of a general theory, which would yield truth *simpliciter* as another special case?

The representational understanding of model theory seems to yield the positive answer to this question. If we take models to be representations of the possible states of our world, then, as we have seen in Chapter 4, model theory turns out to offer not only an account for necessary truth, but also, and primarily, an account for truth *simpliciter*, by means of accounting for the way the world makes statements true or false.

This seems to be the core of Tarski's intensively discussed correspondence theory of truth; and it seems to be also in accordance with the traditional approach to the concept of truth. According to Aristotle's view quoted by Tarski (1944, p.343)

> To say of what is that it is not, or of what is not that it is, is false, while to say of what is that it is, or of what is not that it is not, is true.

In Tarski's own words (ibid., p.343)

> A sentence is true if it designates an existing state of affairs.

Hence (ibid., p.362)

> The sentence "snow is white" is true if, and only if, snow is white.

The sense of such "T-sentences" seems to be quite straightforward: they appear to articulate the world's way of "wielding its influence" on statements. This particular T-sentence says that a statement is true just in case the world is such that snow is white in it. This indicates that what Tarski was after was a theory of *relational* truth; that T-sentences aim at defining a binary relation TRUE-WRT taking a statement as the first argument and (a representation of) a world as the second. The T-sentence *the sentence "snow is white" is true if, and only if, snow is white* states that

(*snow is white* TRUE-WRT w) iff snow is white in w

Such a relational truth-predicate seems to be the required general basis for defining more specific cases, notably the predicate NT of being necessarily true and the predicate T of being true: $s \epsilon NT$ can be defined as $\forall w.(s$ TRUE-WRT $w)$, and $s \epsilon T$ as (s TRUE-WRT w_{act}), where w_{act} is (the representation of) the actual world. The theory of necessary truth as well as the theory of truth *simpliciter* thus seem to be two specific cases of a general theory of relative truth: the general problem is the truth of a statement w.r.t. a given class C of worlds or model structures; the problem of necessary truth we then obtain if we identify C with the class of all possible structures, that of truth *simpliciter* if we identify it with the class containing just the model structure representing the real world.[67]

6.2 Relative Truth

However, we have concluded that the representational understanding of model theory is untenable. Model structures as posited by model theory are auxiliaries of the model-theoretic account for necessary truth, not independent representations of worlds or states of the world. To say that necessary truth is the truth w.r.t. every model-structure is not to say that truth in such-and-such

[67]Tarski (1956, p. 199) puts it as follows: "the concept of *correct or true sentence in an individual domain* [is] a concept of a relative character [that] plays a much greater part than the absolute concept of truth and includes it as a special case". For a discussion see Davidson (1973).

a situation would be truth w.r.t. some more specific class of structures; the identification of necessary truth with universal satisfiability is *constitutive* to the space of model structures or possible worlds. This means that equating s∈NT with ∀w.(s TRUE-WRT w) cannot be understood as the definition of NT, but rather as an implicit definition of TRUE-WRT and the domain alleged to the variable w. This domain is constituted in that we decide to perceive NT as the class of *general* truths, it is the domain of the alleged generality.

Let us imagine that we are to specify a subclass C^* of a class C. One of the possible ways we can do this is to give a collection $<C_i>_{i \in I}$ subclasses of C such that their intersection yields C^*.[68] As to give a collection of sets is clearly the same as to give the collection of their characteristic functions, giving $<C_i>_{i \in I}$ is the same as giving a collection $<f_i>_{i \in I}$ of functions from C to {0,1} such that $x \in C^*$ iff $f_i(x)=1$ for every $i \in I$; and this can in turn be clearly seen to be the same as specifying a relation R between the elements of C and those of I such that $x \in C^*$ iff ∀y.(x R y).[69] In our specific case, C can be seen as the class of all statements, C^* as the class of necessary truths, C_i's as admissible classes of truths (consistent theories), f_i's as admissible truth valuations, i's as possible worlds or model structures, and R as the relation of being true w.r.t. a possible world. However, the procedure described makes a nontrivial sense only in cases where we are able to specify $<C_i>_{i \in I}$ independently of specifying C^*, and this is what we, in our particular case of necessary truth, cannot do - for we cannot know what might be possibly true without knowing what is necessarily true.

If logic employs interpretations and model structures, then these interpretations are to be understood as abstract means of accounting for necessary truth; truth in a model structure should *not* yield truth *simpliciter*, and if it is treated as if it should, the results are absurd. To clarify this, let us first consider the simplest model-theoretic tool, a truth table. Let us assume that on the left side of the table we have the truth values of the statements *Aristotle is human* and *Plato is mortal*, and on the right side the corresponding truth values of the statement *Aristotle is human and Plato is mortal* (Table 10). How is such

---

[68]The reason might be to thereby represent an "incomprehensible" (e.g. non-recursive) class in terms of comprehensible (recursive) classes.

[69]See (30') of Section 5.2.

a table to be understood? Is it not a depiction of the truth conditions of the statements on the right side?

Aristotle is human	Plato is mortal	Aristotle is human and Plato is mortal
T	T	T
T	F	F
F	T	F
F	F	F

Table 10

One way such a table can obviously be understood is as describing the dependence of the truth value of the statement on the right on the truth values of those on the left. Our particular truth-table tells us, for example, that whenever *Aristotle is human* and *Plato is mortal* both are true, *Aristotle is human and Plato is mortal* is *eo ipso* also true. It tells us that whenever either *Aristotle is human* or *Plato is mortal* is not true, *Aristotle is human and Plato is mortal* is *eo ipso* not true.

Is this way of viewing truth tables also applicable to more intricate tools of model-theoretic analysis? Let us imagine that what is on the left side of the table are not truth values of *Aristotle is human* and *Plato is mortal*, but rather some kinds of models or possible worlds. In order to be able to make sense of such a "truth-table", it must be clear for each of the models on the left side which statements are true w.r.t. it; in our particular case it must be clear whether *Aristotle is human* and *Plato is mortal* are true w.r.t. it. If we are to be able to consider the truth value of *Aristotle is human and Plato is mortal* with respect to a world, then we simply *must* know something about this world, in particular we must know if Aristotle is human and Plato is mortal in that world.

This indicates that the primary perspective in which we can view truth-tables, but also other tools of model theory, is the perspective which shows them as summarizing the dependence of truth-values of a statement on truth values of some other statements. Let us call this perspective *truth-dependential.* Now to say, as we do, that the truth-dependential perspective is the primary one is not to say that there would not be other perspectives, such as a "representational" one, or that the other perspectives could not in some circumstances be natural or plausible. (We can *always* say "'x' is true when y" or "'x' is true in a world in which y" instead of "'x' is true when 'y' is true". Thus instead of "'*Aristotle is human and Plato is mortal*' is true when '*Aristotle is human*' is true and '*Plato is mortal*' is true" we may say "'*Aristotle is human and Plato is mortal*' is true when Aristotle is human and Plato is mortal" or "'*Aristotle is human and Plato is mortal*' is true in a world in which Aristotle is human and Plato is mortal".) However, every talk based on such an alternative perspective can be simply translated into the truth-dependential talk.

It is evident that any explicit, theoretical account of the influence of the world on the truth value of a statement 'y' must take the form of "'y' is true when x", which is the same as "'y' is true when 'x' is true". This holds also for cases such as "'y' is true in the world w", for in order to be able to understand this as an account of the influence of the world on the truth value of 'y' we must in such way or another describe w, and this means that we must give some 'x' such that "'y' is true in w" becomes expressible as "'y' is true when x"[70]. This holds also if there is no "canonical" description of w: 'x' could then be the disjunction of all its acceptable descriptions.

On the other hand, and this is crucial, not every truth-dependential talk is meaningfully translatable into the representational talk. To illustrate this, let us consider Table 10. Its first row tells us that if Aristotle is human and Plato is mortal, then *Aristotle is human and Plato is mortal* is true. Let us, however, consider the truth table in which the statement *Aristotle is human and Plato is mortal* is both on the left and on the right side. The first row of this table, viewed representationally, brings us precisely the same information as that of the previous table: that if Aristotle is human and Plato is mortal, then *Aristotle is human and Plato is mortal* is true. However, the former table tells us something nontrivial, while the latter does not. This fact, perfectly clear from the truth-dependential perspective (the former table exposes the dependence of

[70]A statement of the form "there is a world w such that 'y' is true in w" is, of course, meaningful without any description of w, but such a statement is not a description of the influence of the world on the truth value of 'y', it is only a way of saying that 'y' is not necessarily false.

the truth value of a statement on truth values of *other* statements, while the latter does not), cannot be accounted for from the representational perspective. This means that although every representational talk is translatable into a truth-dependential talk, the inverse does not hold: not every truth-dependential talk is meaningfully translatable into the representational talk; hence it is the truth-dependential perspective which is primary.

We can describe the conditions under which a statement is true, but only by an appeal to some other statement describing the conditions. But this is only a relative description: we take the truth conditions of the describing statement for granted. There is no way of an *absolute* presentation of the conditions under which a statement is true. As Wittgenstein (1977, p.27) puts it

> Die Grenze der Sprache zeigt sich in der Unmöglichkeit, die Tatsache zu beschreiben, die einem Satz entspricht (seine Übersetzung ist), ohne eben den Satz zu wiederholen.[71]

What we must conclude is: *there is no account of how truth values of statements are influenced by the world beyond the account of how truth values of statements are influenced by truth values of other statements*. (In other words: there is no representational semantics beyond truth-dependential semantics.) However, on the other hand, there are aspects of semantics which can be understood *only* from the truth-dependential perspective (and not from the representational one); so there *is* truth-dependential semantics beyond representational semantics.

We are not able to say anything nontrivial about the truth of an individual statement; hence the only thing we can do is to theorize about relative truth. However, what we are able to do is to theorize about truth of a statement relative not to the state of the world, but rather to the truth of other statements; thus what we are doing is theorizing about truth-dependence. But truth-dependence is nothing else than consequence; hence in this sense, there is no theory of truth beyond the theory of consequence.

[71]The limit of language is shown by it being impossible to describe the fact which corresponds to (is the translation of) a sentence, without simply repeating the sentence.

As an example of a standpoint contrary to the one adopted here, let us discuss the claim of Bealer (1993, p.12)[72]:

> I simply want to know in general what conditions must hold for an arbitrary atomic intensional sentence in an arbitrary language to be true, rather than not true. The question is clear; it would be mysterious in the extreme if it did not have an answer. I shall adopt the premise that a *referential* theory provides the only viable answer to this question. On a referential theory, an atomic intensional sentence $\ulcorner F[A] \urcorner$ is true if and only if there is something that the singular term $\ulcorner [A] \urcorner$ designates and the predicate $\ulcorner F \urcorner$ applies to that thing.

The Wittgensteinian objection to this premise is the following: what if we now want to know in general what conditions must hold for an arbitrary predicate to apply to the thing designated by an arbitrary singular term? This is evidently a question no easier to answer, but also no less clear, than the one posed by Bealer, and if we accept his line of reasoning, then we should expect it to have an answer. However, it seems to be obvious that every new answer gives birth to a new clear question, and sooner or later we have to stop answering; hence that there we will necessarily be left with a clear question that remains unanswered.

Note that this is not to say that a referential theory of the kind urged by Bealer would be of no use. We have seen in the previous chapter that such a theory provides a way to account for certain necessary truths and certain patterns of inference and hence to envisage certain aspects of the inferential potential of our language. However, what it does *not* provide for is precisely that for which Bealer invokes it, namely the answer to the question when is an individual statement true. In other words, the theory can help us see how truth values of statements depend on those of other statements, but it cannot explain, over and above this, how truth values of statements depend on the world.

The Tarskian approach to truth is controversial: it is surely a theory of truth, because it says us something about truth; but at the same time, it is not an explication of truth, because it does not tell us what truth is - for in

---

[72]Bealer speaks about *intensional* statements, but his referential theory is evidently general.

explicating truth for a given language it "necessarily failed to specify how to go on to further cases" (Davidson, 1990, p.287). Now we can see that this is the case because the theory explicates truth only in the sense of listing all truths (which is a nontrivial enterprise due to the infinite number of truths in any language worth its name). What the approach really does explicate is *truth-dependence*.

6.3 The Correspondence Theory

The explication of the concept of truth implied by Tarski's correspondence theory of truth can be articulated as follows:[73]

> The truth of a sentence consists in its agreement with (or correspondence to) reality.

The correspondence theory is undoubtedly appealing. It accords with the commonsense view of truth and also does not appear to contradict any deeper insight into the nature of truth. However, is this theory really capable of *explaining* truth? We know with certainty that many statements of our language are true. Does this mean that we know that they correspond to reality *prior to the knowledge of their truth* or *over and above the knowledge*? What should the nature of such knowledge be? Should it mean that we know that reality is such-and-such and that the statement is in its way also such-and-such? This would presuppose that we have an immediate access to reality, bypassing language. But if we accessed reality in such an immediate way, what would its such-and-suchness mean? Could it be commensurate with the such-and-suchness of language?

The Tarskian solution amounts to expressing the such-and-suchness of reality by a statement of some language. Of course of a language different from the language which is the object of investigation, of a *meta*language. Metalanguage plays the role of an "immediate" presentation of reality. Such a picture may be useful for some purposes; it is, however, nothing more than a

[73]Tarski (1944, p.343).

picture[74] which is hardly capable of *explaining* truth. To address truth as such we cannot restrict ourselves to languages which can be encompassed in a metalanguage, we must be able to address language in its universality. We may reduce truth in one language to truth in another language; however, we can hardly indulge in this *ad infinitum*, and from the point of explication of the concept of truth it is no more than a vicious circle.[75]

So how to explain truth? To answer this, we must first ask other questions: In which sense is truth capable of being explained? And in which sense is it in a need of being explained? Tarskian T-sentences can be understood to explain truth in the sense that they reduce metalinguistic judgements, including the predicate *is true* to statements containing no such predicate. That this is possible is an important (as well as trivial) fact; however, it neither enables us to dispense with the predicate *is true* altogether[76], nor does it tell us *why* a statement is true.

To explicate means to "translate" into some terms which are in a sense more basic or more primitive; but if we do not want to run in a circle, at some point we must stop. If we succeed in explicating one concept, we must pay for it by giving up the explication of some other concepts. But if this is so, is it not just *truth* that should play the role of the basic unexplicated notion? As Davidson (ibid., p.314) puts it, "truth is one of the clearest and most basic concepts we have, so it is fruitless to dream of eliminating it in favor of something simpler or more fundamental."

---

[74]In Wittgenstein's sense.

[75]As Romanos (1983) comments on Quine's criticism of Carnap: "Saying exactly, in a noncircular way, what it means for some sentence to be true - explaining, so to speak, the basis or ground of its truth - will ordinarily presuppose understanding some other sentence of which the former may be seen as translation. The truth of those two sentences will never derive from the fact that they are translation of one another, though it will explain why they are."

[76]We may say *Snow is white* instead of *'Snow is white' is true*, but there is no such paraphrase for *He wanted to say something that is true*.

6.4 The Davidsonian Approach to Truth

Does this Davidsonian stance imply that there is no sense at all in attempting at a theory of truth? No - to say that we should not attempt at explicating truth in the sense of reducing it to a simpler concept does not mean to say that there is nothing theoretical to be said about truth. Davidson himself, despite his reservation about the possibility and reasonability of trying to eliminate the concept of truth, repeatedly stresses the importance of a theory of truth, claiming that it is the theory of truth that is capable of facilitating our understanding of semantics of language.

What is the character of the Davidsonian approach to the theory of truth? Davidson's (1973, pp.69-71) basic point is that "there are important differences between theories of relative, and of absolute, truth, and the difference makes theories of the two sorts appropriate as answers to different questions". We do not need, and we in fact cannot have, a theory of truth which would account for the relationship between statements and the world, we need merely a finite account for the infinity of T-sentences. Such an account has to "work by selecting a finite number of truth-relevant expressions and a finite number of truth-affecting constructions from which all sentences are composed" and by way of providing "an analysis of structure relevant to truth and to inference" it gives us the key to such concepts as logical form and meaning.

In previous chapters we saw that natural, pre-theoretical concepts can be accounted for via criterial reconstruction; we saw that grammar can be seen as aiming at a criterial reconstruction of well-formedness, and logic at a criterial reconstruction of consequencehood or necessary truth. We have stressed that logic can be made sense of only if we do not consider it as aiming at the elimination of necessary truth, but rather only at its reconstruction. Now we can say that truth can be, analogously, accounted for via criterial reconstruction of the class of true statements. And such a criterial reconstruction seems to be precisely what Davidson has been urging: selecting a finite number of basic expressions, a finite number of rules and reconstructing the class of true statements as recursively built by these means.

However, let us stress the difference between reconstructing necessary truth and reconstructing truth *simpliciter*. To reconstruct we have to know (although pre-theoretically) what is to be reconstructed. We have stressed that necessary statements are such that their truth values can be determined without "investigating into the world"; so the reconstruction of necessary truths is an armchair enterprise (which in no sense means that it is trivial) and hence it is a fair game for a logician or a philosopher. To find out about truth *simpliciter*, on the other hand, is in general the matter of natural science; and hence it is a task neither for a logician, nor for a philosopher. What a logician or a philosopher can do is to provide pieces of such a reconstruction, pieces concerning the eternal, world-independent aspects of truth. Her or his business is not to account for the truth of the fact that the statement *Aristotle is mortal* is true, but it *is* her or his business to account for the fact that it is true whenever *Aristotle is human* and *Every human is mortal* are. In other words, she or he can account for truth only through accounting for consequence. Hence: the aspect of truth a philosopher can theorize about comes out as a theory of consequence, or of necessary truth. Tarskian correspondence theory invokes the illusion that we can account for truth beyond accounting for necessary truth; that we can make a theory of the world's wielding its influence on statements. However, we have seen that there is no such theory beyond the theory of consequence.

We have stated that criterial reconstruction works toward restating an infinite range of items as generated by recursive application of a finite collection of rules. Such criterial reconstruction can be seen as imposing a certain structure on items of the range; the structure of an item being determined by the way the item gets generated. In the case of the criterial reconstruction of the range of true statements the imposed structure is particularly important: it leads us to seeing truth in terms of reference and meaning. We shall see this in the following chapters.

6.5 The Theory of Truth as Its Criterial Reconstruction

To illustrate the basic difference between the representational, approach to truth and our, Davidsonian, approach, let us look at the role Tarskian T-sentences can play within our theory.

Let us recall the way we understood the theoretical account for necessary truth. We assumed that there is a pre-theoretical, uncriterial, class NT of necessary truths and we tried to define a criterial class to coincide with it and thus to be understood as a criterial reconstruction of NT. The criterial class might be the class of theorems of an axiomatic system, the class of statements that are valid, or something else; anyway, if we propose a class NT^* as the theoretical reconstruction of NT, then our theory is adequate if for every statement s

$s \epsilon NT^*$ if and only if $s \epsilon NT$

The same now holds for a theoretical reconstruction of truth: we may propose various ways of reconstructing the pre-theoretical, uncriterial - and therefore "incomprehensible" - class T of true statements by means of a criterial - "comprehensible" - class T^*; but the condition of adequacy of any such theory would be that for every statement s

$s \epsilon T^*$ if and only if $s \epsilon T$

Expanding '$s \epsilon T$' to its full wording we have

$s \epsilon T^*$ if and only if s is true

Let now 'p' be the statement such that s='p'; we have

'p'ϵT^* if and only if 'p' is true

and as 'p' is true if and only if p, we further have

'p'ϵT^* if and only if p

We can trivially alter our theory by employing the predicate true^* instead of the class T^* (writing 's is true^*' in place of '$s \epsilon T^*$'); and if we believe that misunderstanding is not possible, we can even drop the asterisk and use simply 'true' instead of 'true^*' for our theoretical predicate. In this way we reach

'p' is true if and only if p

which is nothing else than Tarski's T-sentence. However, we take this sentence as amounting not to a correspondence between language and the world, but as expressing an adequacy condition for our theory. Moreover, the sentence tells us nothing about the natural, pre-theoretical concept of truth (let alone how to reduce it to other concepts), it concerns what aims at being a theoretical reconstruction of the concept[77].

It seems then, as Davidson (1990) argues (following Soames, 1984, Putnam, 1985, Etchemendy, 1988, and others) that the Tarskian formal definition of the truth predicate cannot be the whole story we are able to tell about truth. Of course, a formally defined concept cannot by itself tell us anything about a natural concept. Therefore, the whole story is the Tarskian formal theory plus its projection onto a pre-theoretical language with its truths.[78] However, in contrast to Davidson (ibid., p.292) I do not think that this could be improved by enhancing Tarski's theory with a "truth axiom" stating that the formal truth-predicate is coextensional with its natural counterpart. Any such enhancement would lead to a further formal theory, and this theory, in order to be understandable as a theory of something, would again have to be projected onto the informal reality.[79]

---

[77]However, as in this way we manage to state the adequacy condition for our theoretical predicate 'true*' without invoking the natural predicate 'true', there appears to be the possibility - perverse, indeed - of analogous "adequacy conditions" for the natural predicate itself. However, we can hardly make sense of the natural predicate's being inadequate, hence we seem to be left with understanding the conditions as *characterizing* truth - the way in which the concept of truth is explicated by Horwich (1990).

[78]As Stekeler-Weithofer (1992) points out, this way of attacking problems and puzzles of our world goes back to Plato - we posit formal, ideal objects ("ideas", "models", "formal theories") which we then project back onto the real world ("methexis", "interpretation", "instantiation").

[79]The same reservation can be expressed about the whole Davidsonian program of employing Tarski's formal theory for the purposes of explicating the concept of truth: it is - at least partly - an attempt to formalize the adequacy conditions for the projection of a formal theory onto that which the theory purports to be the theory of; and this can yield nothing else than another formal theory in the same need of projection. It is also worth noting that, as pointed out by Kripke (1976, p.405), this program is nontrivial only if we do not allow for the really substitutional quantification. The "really substitutional" quantification would (putting things more radically than Kripke) render the Tarskian characterization of truth trivial because it would allow us to quantify into quotational contexts (more precisely: into what we see as quotational contexts from the viewpoint of referential quantification, for a really substitutional quantifier views its scope as a mere sequence of signs) and so to capture the whole theory of truth by a single statement, namely "**Π**p.'p' is true iff p" - meaning that whatever we substitute for the letter "p" into the sign sequence "'⌢p⌢'⌢is⌢true⌢iff⌢p" we get a sign sequence which is a true statement. Note that this is

6.6 The Primacy of Truth

A concise articulation of the standpoint following from the above considerations has been given by Austin (1961, pp.123-124):

> When a statement is true, there is, *of course*, a state of affairs which makes it true and which is *toto mundo* distinct from the true statement about it: but equally of course, we can only *describe* that state of affairs *in words* (either the same, or, with luck, others). I can only describe the situation in which it is true to say that I am feeling sick by saying that it is one in which I am feeling sick (or experiencing sensations of nausea): yet between stating, however truly, that I am feeling sick and feeling sick there is a great gulf fixed.

The fact pointed out by Austin is so simple that it may seem banal; it is, however, a fact which is fatal to any attempt at an explicit correspondence theory and at a representational semantics. This fact is the recognition which means the abandonment of semantic naivety; it is one of the facts which caused Wittgenstein to abandon his "crystal clear" system of the *Tractatus* in favour of his later disordered considerations.

That correspondence cannot be addressed explicitly and that it, therefore, cannot help explicate truth has been recognized by many theoreticians[80]. Such recognition has usually led to embracing relativism and the coherence theory of

---

precisely what we *can* do in natural language - the statement "for every sentence p, 'p' is true iff p" will indeed be understood in this "substitutional" way.

[80]As early as 1935 Hempel states the limitation of the theory quite clearly: "None of those who support a cleavage between statements and reality is able to give a precise account of how a comparison between statements and facts may possibly be accomplished, and how we may possibly ascertain the structure of facts." (Hempel, 1935, p.51). Twenty four years later, Dummett could declare the theory of correspondence as wholly *passé*: "... we have nowadays abandoned the correspondence theory of truth; and justify our doing so on the score that it was an attempt to state a *criterion* of truth in the sense in which this cannot be done." (Dummett, 1959, p. 14). However, it would be wrong to read this declaration as an epitaph of the correspondence theory - the theory continues to play an influential role even nowadays, albeit wearing ever more sophisticated guises.

truth. However, if we concede that there is no necessity of a theory of truth (with the exception of Horwich's, 1990, minimal theory restricted to putting the equal sign between stating that a statement is true and the statement itself; and with the exception of a theory of truth as a systematization of truths), then there is actually no need to ban the correspondence view of truth. The only change required is that it is no longer possible to consider correspondence as a *theory*, i.e. as something which may be found out to be true or false. Indeed, the above considerations do not suggest any evidence *against* correspondence, rather they suggest evidence *against the verifiability and falsifiability* of correspondence - against the meaningfullness of understanding correspondence as something over and above truth, let alone as a *criterion* of truth.

This is to say that we need not reject the statements *'p' is true if and only if p* in the sense *'p' is true if and only if it is a fact that p*; we only have to reject that they could be taken as contingent theses. If someone thinks that he needs the theoretical concept of fact, then such statements can be well used to constitute its implicit definitions. Similarly we can take statements *'p(t)' is true if and only if the individual t instantiates the property p* as constitutive to the concepts of individual, property and instantiation.

So it is not necessary to reject the correspondence view, it is only necessary to understand it differently. If we accept that we have no approach to things-in-themselves to compare them with expressions[81], then we have to accept that there is essentially no way to verify correspondence[82]; and that implies that there is no theory of correspondence as an empirical theory about the world. However, it also implies that there is no way to *falsify* correspondence, and hence nothing can prevent us from *stipulating* correspondence - if we believe that such stipulation is of some use.[83]

This means that not an agreement of reality with language verifies correspondence, but rather that our adherence to correspondence "constitutes"

---

[81]No *innocent eye*, as Goodman (1960) puts it.

[82]And note that this does not mean that we could not find out whether a statement corresponds to reality, it rather means that there is no such thing to be found out over and above the finding out of whether the statement is true.

[83]And if we beware falling into the illusion that by stipulating correspondence we have explained the concept of truth or the working of language. For all in all, correspondence is, as Quine (1994, p.496) puts it, "an idle positing of entities solely to create correspondence. It is pernicious, engendering an illusion of explanation."

reality as agreeing with language. Not: *Aristotle is human* is true *because* the individual Aristotle has the property of being human; but rather: that *Aristotle is human* is true can be understood as Aristotle's being human. Or, if we want to sound more Wittgensteinian, not: *Aristotle is human* says that Aristotle has the property of being human; but rather: *Aristotle* concatenated with *to be human* says that Aristotle is human.[84] This is to say that *the idea of correspondence is constitutive to ontology*. In this sense, we may well join Leilich (1983, p.22) in speaking about a new 'Copernican turn': "So wie im Denken des Mittelalters die Welt sich nach dem intellectus divinus richten muß, so muß eine Welt, die für uns begreiflich sein soll, sich nach unseren Sprachregeln richten."[85]

[84]See footnote 61.

[85]"Just like the Middle Age conception that the world must obey the intellectus divinus, the world, if it is to be comprehensible to us, must obey the rules of our language."

7 Denotation

7.1 From Satisfaction to Denotation

We saw in Chapter 5 that model theory, in accounting for the traditional range of necessary truths (i.e. for those addressed by classical logic), yields a framework of individuals and properties, which enables us to accommodate the pre-theoretical concepts of reference and existence. The existence of the individuals and properties was not the result of a metaphysical deliberation, but resulted from our way of reconstruction of the inferential structure of language.

Now the idea is that if the same approach is applied to a wider range of necessary truths, then what results is a framework which instead of being suited to explicate the concept of reference, is suited rather to explicate the general concept of meaning; hence that the difference between reference and meaning is only "quantitative", not "qualitative". To move from reference to meaning we have to do two things: first, we must stop treating the difference between extralogical and logical expressions as the difference between expressions which do refer and those which do not; we must consider *all* expressions as acquiring "semantic values". Second, we must extend the range of necessary truths (i.e. patterns of consequence) accounted for beyond those that gave rise to classical logic. Let us first turn our attention to the former theme.

Tarski's approach was based on the sharp distinction between logical and extralogical constants; the extralogical constants were the only ones which were subject to reference-assignment. What would it mean to treat *all* expressions as referring? Well, it need not mean very much. We can simply consider each logical constant as referring to a fixed object - there is clearly no substantial difference between saying that a constant's reference is fixed and saying that the constant does not refer at all; the point of the logical/extralogical distinction is the difference between the fixed and the alternating, which is retained.

Then we can make a further move: we can extend reference to complex expressions. To every grammatical rule, we give a rule computing the reference of the output of the rule from the reference of the input. This can also be done without really substantially modifying the original framework: we can simply turn the recursive constraints characterizing the relation of satisfaction into rules of composition of reference.

Let us first consider elementary statements. We saw that satisfaction can be considered as a relation between a statement and a reference-assignment, but instead of saying that an elementary statement $P(T_1,\dots,T_n)$ is satisfied by a reference-assignment r, we can clearly also say that $P(T_1,\dots,T_n)$ is satisfied by the objects $\mathrm{r}(P),\mathrm{r}(T_1),\dots,\mathrm{r}(T_n)$, so we can consider satisfaction as a matter of comparing $P(T_1,\dots,T_n)$ with an (n+1)-tuple of objects. Now we can consider this comparison as consisting of two steps: we can consider the (n+1)-tuple of objects to first add up to a single "summary" object and this "summary" object to be then compared with the statement. Thus, instead of comparing $\mathrm{r}(P),\mathrm{r}(T_1),\dots,\mathrm{r}(T_n)$ with $P(T_1,\dots,T_n)$, we first let $\mathrm{r}(P),\mathrm{r}(T_1),\dots,\mathrm{r}(T_n)$ yield an object $\mathrm{r}(P(T_1,\dots,T_n))$ and then compare $\mathrm{r}(P(T_1,\dots,T_n))$ with $P(T_1,\dots,T_n)$. This means that we stop treating the relation of satisfaction between statements and predicate-referents plus term-referents as primitive, and start considering it instead as derived from a function composing statement-referents out of the corresponding predicate-referents and term-referents, and a relation of satisfaction between statements and statements-referents. This requires defining (i) functions $\mathrm{F_n}$ (for every natural n) such that $\mathrm{r}(P(T_1,\dots,T_n)) = \mathrm{F_n}(\mathrm{r}(P),\mathrm{r}(T_1),\dots,\mathrm{r}(T_n))$ for every n-ary predicate P and terms $T_1,\dots,T_n$; and (ii) a relation $\mathrm{SAT^S}$ such that $\mathrm{r}(P(T_1,\dots,T_n))\ \mathrm{SAT^S}\ P(T_1,\dots,T_n)$ either holds or not for every elementary statement $P(T_1,\dots,T_n)$. All these changes are insubstantial from the viewpoint of the resulting theory provided we ensure that the satisfaction relation yielded by the combination of $\mathrm{F_n}$ and $\mathrm{SAT^S}$ does not differ from the original relation SAT; that $\mathrm{r}(P),\mathrm{r}(T_1),\dots,\mathrm{r}(T_n)\ \mathrm{SAT}\ P(T_1,\dots,T_n)$ if and only if $\mathrm{F_n}(\mathrm{r}(P),\mathrm{r}(T_1),\dots,\mathrm{r}(T_n))\ \mathrm{SAT^S}\ P(T_1,\dots,T_n)$.

However, at this stage it is reasonable to change the terminology. It is no longer adequate to speak about reference, so hereafter we shall use the term *denotation*. Thus, each reference-assignment gets extended to a denotation-assignment: the two assignments coincide for constants, but the denotation-assignment, unlike reference-assignment, assigns values also to statements. We shall continue calling the objects denoted by terms *individuals* and those denoted by predicates *properties* and *relationships*; the entities denoted by statements will be called *(proto-)facts* (as proposed in Section 4.2). If S is a statement, then instead of saying that the fact $\mathrm{r}(S)$ satisfies (does not satisfy) S we shall say that the fact $\mathrm{r}(S)$ is actual (unactual). (This terminology turns out to be somewhat odd in the case of classical logic, because there we can clearly make do with just two "facts", and it is, of course, natural to take them directly as the two truth values; but it is needed for the more general case.) Remember that this can still be considered as merely a formal manoeuvre: it is enough to ensure that if d is the denotation-assignment which arises out of the reference-assignment r, then d($P(T_1,\ldots,T_n)$) is an actual fact if and only if r satisfies $P(T_1,\ldots,T_n)$.

Now we can continue and extend denotation to all complex expressions. For every grammatical rule $\mathbb{R}$ we must define a rule $\mathbb{R}^*$ such that the denotation of the output of $\mathbb{R}$ equals the value yielded by the application of $\mathbb{R}^*$ to the denotations of its input, i.e. that $\mathrm{d}(\mathbb{R}(E_1,\ldots,E_n)) = \mathbb{R}^*(\mathrm{d}(E_1),\ldots,\mathrm{d}(E_n))$ for every n-tuple $E_1,\ldots,E_n$ of expressions from the domain of $\mathbb{R}$. This can be done in various ways, but again, if we want to keep the whole enterprise as a mere technical variant of the original satisfactional apparatus, we must guarantee that if S is a statement, r a reference-assignment and d the corresponding denotation-assignment, then d(S) is an actual fact just when r satisfies S.

This can be achieved by basing the definitions of the rules of composition of denotation, the companions $\mathbb{R}^*$'s of $\mathbb{R}$'s, on the corresponding satisfaction constraints. Let us take an example. If S_1 and S_2 are statements and o a binary operator (like $\wedge$), then $S_1 o S_2$ is a statement; in other words, we have a grammatical rule - call it $\mathbb{R}_{O2}$ - yielding a statement whenever fed with two statements and a binary operator. The satisfaction constraints applying to the statements that are output from this rule are (see Section 4.4):

$$\mathrm{r\ SAT\ } \mathbb{R}_{O2}(S_1, \wedge, S_2) \text{ iff r SAT } S_1 \text{ and r SAT } S_2 \qquad (\mathrm{C}\wedge)$$
$$\mathrm{r\ SAT\ } \mathbb{R}_{O2}(S_1, \vee, S_2) \text{ iff r SAT } S_1 \text{ or r SAT } S_2 \qquad (\mathrm{C}\vee)$$
$$\mathrm{r\ SAT\ } \mathbb{R}_{O2}(S_1, \rightarrow, S_2) \text{ iff r SAT } S_2 \text{ or not r SAT } S_1 \qquad (\mathrm{C}\rightarrow)$$

We have seen that turning reference into denotation means turning 'r SAT S' into 'd(S) belongs to the class A of actual facts', and this makes it possible to turn these constraints directly into constraints on denotation

$$\mathrm{d}(\mathbb{R}_{O2}(S_1, \wedge, S_2)) \epsilon \mathrm{A} \text{ iff } \mathrm{d}(S_1) \epsilon \mathrm{A} \text{ and } \mathrm{d}(S_2) \epsilon \mathrm{A} \qquad (\mathrm{C}\wedge')$$
$$\mathrm{d}(\mathbb{R}_{O2}(S_1, \vee, S_2)) \epsilon \mathrm{A} \text{ iff } \mathrm{d}(S_1) \epsilon \mathrm{A} \text{ or } \mathrm{d}(S_2) \epsilon \mathrm{A} \qquad (\mathrm{C}\vee')$$
$$\mathrm{d}(\mathbb{R}_{O2}(S_1, \rightarrow, S_2)) \epsilon \mathrm{A} \text{ iff } \mathrm{d}(S_2) \epsilon \mathrm{A} \text{ or not } \mathrm{d}(S_1) \epsilon \mathrm{A} \qquad (\mathrm{C}\rightarrow')$$

Such constraints can be further turned into constraints on the companion $\mathbb{R}_{O2}{}^*$ of $\mathbb{R}_{O2}$, i.e. on the denotation-composing function for which $d(\mathbb{R}_{O2}(S_1,o,S_2))=\mathbb{R}_{O2}{}^*(d(S_1),d(o),d(S_2))$:

$\mathbb{R}_{O2}{}^*(d(S_1),d(\wedge),d(S_2))\epsilon A$ iff $d(S_1)\epsilon A$ and $d(S_2)\epsilon A$ (C∧'')
$\mathbb{R}_{O2}{}^*(d(S_1),d(\vee),d(S_2))\epsilon A$ iff $d(S_1)\epsilon A$ or $d(S_2)\epsilon A$ (C∨'')
$\mathbb{R}_{O2}{}^*(d(S_1),d(\rightarrow),d(S_2))\epsilon A$ iff $d(S_2)\epsilon A$ or not $d(S_1)\epsilon A$ (C→'')

Performing this systematically for all grammatical rules yields an alternative formal apparatus, which, however, supports the same notion of validity and hence leads to the same reconstruction of necessary truth as the original one. The difference is merely in that now we can speak about the semantical value (denotation) of every expression. We retain the logical/extralogical bipolarity as the bipolarity between constants with fixed, and those with freely varying, reference.

7.2 Denotation and Substitution

Prior to the Tarskian definitive satisfactional approach to necessary truth, we were exploring the provisional substitutional approach; let us now return to this point. We have seen that a substitution can be considered as an automorphism of the algebra of language, in our particular case an automorphism induced by a category-preserving mapping of extralogical words on extralogical words. As we saw in Section 4.6, Tarski proposed replacing the mapping of extralogical words on extralogical words by the mapping of extralogical words on some language-external entities ("word-referents").

This prompted a substantial modification of the whole framework. We replaced SUBST by REF; however, while in the substitutional case we defined

$s\epsilon V \equiv_{Def} f(s)\epsilon T$ for every $f\epsilon SUBST$

defining in the satisfactional case analogously

$s\epsilon V \equiv_{Def} r(s)\epsilon T$ for every $r\epsilon REF$

would make no sense - r is not applicable to s, for reference-assignments act on words only. The problem is that the assignment of elementary extralinguistic entities to elementary extralogical expressions does not by itself induce the assignment of extralinguistic objects to complex expressions; and even if we manage to get the assignment extended to complex expressions and to statements, the entities assigned to statements would not be themselves statements and hence would be neither true, nor false. The way proposed by Tarski and accepted by us in the previous chapters was to introduce an unanalyzed relation SAT which for every statement s and every reference-assignment r determines whether the former satisfies the latter; then we can define

$$s \in V \equiv_{Def} r \text{ SAT } s \text{ for every } r \in REF.$$

However, the switch from reference to denotation now allows our approach to become more analogous to the substitutional case. The structure of denotations parallels the structure of language, hence denotations can be considered to form an algebra similar to the algebra of language. If the carrier of the language-algebra consists of the grammatical categories $\langle C_i \rangle_{i \in I}$, then the carrier of this new, denotational algebra consists of some domains $\langle D_i \rangle_{i \in I}$. We further assume that the objects in the domain corresponding to the category of statements, (proto-)facts, are divided into two groups, into the group of actual facts and that of unactual ones.

Unlike REF, DEN can be considered as directly analogous to SUBST: whereas SUBST is a class of automorphisms of the language-algebra, DEN is a class of homomorphisms from the language-algebra into the denotational algebra. SUBST can thus be seen simply as a special case of DEN for the case when the denotational algebra is identical with the language-algebra. This means that the new, denotational approach can be considered as a direct way to generalize the original, substitutional approach. If we use A to denote the class of actual facts, then we can define

$$s \in V \equiv_{Def} d(s) \in A \text{ for every } d \in DEN.$$

Besides this possibility of maintaining a closer analogy with the prototypical substitutional case, the switch from satisfaction to denotation introduces a new possibility: we can cease treating the logical and the extralogical as two alternatives such that *tertium non datur*, and start instead to see them as two poles of a continuous scale. This may help us counteract difficulties posed by statements such as $\exists x \exists y(x \neq y)$ considered by Etchemendy (1988 and 1990): all the constants this statement consists of are logical, but the statement appears to be contingent.

Let us call a class H of homomorphisms from the language algebra *fixed w.r.t. an expression e* if there exists a d such that $h(e)=d$ for every $h \in H$; let us call it *free w.r.t. e* if for every d from the appropriate domain and for every $h \in H$, the function h', such that $h'(e)=d$ and $h'(x)=h(x)$ for every other expression x, belongs to H. To say that every constant is either purely logical, or purely extralogical is to say that every constant is either bound to denote one and the same entity, or can freely denote whichever entity of the appropriate domain; hence it is to say that whichever constant we take, the class of denotation-assignments will be either fixed, or free, w.r.t. it.

However, there is now also the possibility of constants that are neither purely logical, nor purely extralogical, constants whose denotation does not vary altogether freely, but only within some bounds. Let us consider the statement $\exists x \exists y \neg(x=y)$. It consists of constants which are usually considered as logical, and whose denotation thus should be fixed. This means that also the denotation of the whole statement should be fixed - which seems to contradict the intuition that the statement expresses a contingent claim, namely that there are at least two individuals. We can hardly consider any of the constants constituting the statement as extralogical - this would deprive us of the possibility of accounting for the necessary truth of many other statements. However, if we stop restricting ourselves to pure logicality and pure extralogicality, then we can explain the situation quite straightforwardly: some of the constants involved in the statement, namely the quantifiers, are not purely logical (although they are far from being purely extralogical). The point is that although a quantifier cannot denote *every* class of classes of individuals, it does not rigidly denote one and the same class of classes of individuals. What we usually perceive as the variation of the universe of individuals can in fact be understood as a variation of the denotations of quantifiers (as far as terms and predicates are concerned, there is no reason not to consider their denotations as being drawn from one all-encompassing universe).[86]

7.3 The Denotational View of Classical Logic

Let us, for the sake of illustration, consider the case of classical logic, i.e. of the classical first-order predicate calculus. We may consider this calculus to be based on the following categories of expressions: S (the category of statements), O^1 (the category of unary operators, consisting of the single constant $\neg$), O^2 (the category of binary operators, consisting of the three constants $\wedge$, $\vee$, $\rightarrow$), P^n (the category of n-ary predicates) for every natural n, T (the category of terms) and Q (the category of quantifiers, consisting of the constants $\forall$ and $\exists$).

The expressions of the categories T, P^n and S are usually considered as referring, namely to individuals, relations among individuals, and truth values, respectively; whereas those of the other categories are taken as non-referring, their semantics being captured by way of axioms. However, it is well-known that this way of treating the predicate calculus is not essential. We may well consider the semantics of *all* expressions as being given via denotations: operators can be thought about as denoting truth-functions and quantifiers classes of classes of individuals. This is the way we view predicate calculus, e.g., if we consider it as embedded within the calculus of the theory of types.[87]

The grammar of the calculus can then be articulated - in a way suitable for our purposes - by means of the rules with the following domains and ranges (R: $A \times B \longrightarrow C$ means that the rule R maps elements of the Cartesian product

[86]This offers a way of understanding the basic change undergone by model theory from the initial Tarski's proposal to its fully-fledged form (which was later embraced by Tarski himself). Tarski's proposal seemed to amount to one definite universe (somehow grounded in the actual world), whereas the fully-fledged model theory is the matter of all possible universes. We can now see this turn as a way of reaching the faithful model-theoretical acount for quantifiers without abandoning the satisfactional account with its logical/extralogical bipolarity.

[87]See Kemeny (1948) or Henkin (1950).

of the sets A and B, i.e. pairs of expressions of the respective categories A and B, on expressions of the category C:

$\mathbb{R}_{O1}$: $O^1 \times S \longrightarrow S$
Example: $\mathbb{R}_{O1}(\neg,\mathbf{Hu(Ar)}) = \neg\mathbf{Hu(Ar)}$

$\mathbb{R}_{O2}$: $O^2 \times S \times S \longrightarrow S$
Example: $\mathbb{R}_{O2}(\wedge,\mathbf{Hu(Ar)},\mathbf{Mo(Ar)}) = \mathbf{Hu(Ar)} \wedge \mathbf{Mo(Ar)}$

$\mathbb{R}_{Pn}$: $P^n \times T \times \ldots \times T \longrightarrow S$
Example: $\mathbb{R}_{P1}(\mathbf{Hu},\mathbf{Ar}) = \mathbf{Hu(Ar)}$

$\mathbb{R}_Q$: $Q \times P \longrightarrow S$
Examples: $\mathbb{R}_Q(\forall,\mathbf{Hu}) = \forall x\mathbf{Hu}(x)$
$\mathbb{R}_Q(\forall,\lambda x(\mathbf{Hu}(x)\rightarrow\mathbf{Mo}(x))) = \forall x(\mathbf{Hu}(x)\rightarrow\mathbf{Mo}(x))$

$\mathbb{R}_\lambda$: $S \times T \longrightarrow P^1$
Example: $\mathbb{R}_\lambda(\mathbf{Hu(Ar)}\rightarrow\mathbf{Mo(Ar)},\mathbf{Ar}) = \lambda x(\mathbf{Hu}(x)\rightarrow\mathbf{Mo}(x))$

The denotation-assignment for the predicate calculus is a function mapping every expression of the language of the calculus on a denotation. Of course, not every such function counts as a denotation: expressions must be assigned objects of appropriate kinds, and some expressions must, moreover, be assigned definite appropriate objects. Thus, an assignment of objects to expressions which is to count as denotation-assignment is bound to assign, e.g., the truth-value switching function to $\neg$, and the class of all non-empty classes of individuals to $\exists$; and similarly for other constants.

In general, the concept of denotation-assignment for the predicate calculus can be defined using the following three types of constraints:

(i) constraints specifying what kinds of objects are to be assigned to expressions of individual grammatical categories;

(ii) constraints specifying which particular objects are to be assigned to certain particular constants; and

(iii) constraints specifying how the denotation of complex expressions is determined by the denotation of their parts.

The constraints of type (i) state that a statement is assigned a truth value; a unary operator is assigned a unary function from truth values to truth values;

a binary operator is assigned a binary function from truth values to truth values; a term is assigned an individual (an element of a given class - "the universe of discourse"); an n-ary predicate is assigned an n-ary function from individuals to truth values (i.e. an n-ary relation among individuals); and a quantifier is assigned a class of classes of individuals. If we write $[A \longrightarrow B]$ for the class of functions from the class A to the class B, then we can articulate these constraints as definitions of the domains corresponding to individual grammatical categories (where U is the universe of discourse):

$$
\begin{aligned}
D_S &= \{T,F\} \\
D_{O1} &= [\{T,F\} \longrightarrow \{T,F\}] \\
D_{O2} &= [\{T,F\} \times \{T,F\} \longrightarrow \{T,F\}] \\
D_T &= U \\
D_{Pn} &= [U^n \longrightarrow \{T,F\}] \\
D_Q &= [[U \longrightarrow \{T,F\}] \longrightarrow \{T,F\}]
\end{aligned}
$$

The constraints of the type (ii) then stipulate the assignment of the appropriate concrete functions to $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\forall$ and $\exists$; they can be formulated, e.g., in the following way:

$$
\begin{aligned}
&d(\neg)(x)=T \text{ iff } x=F \\
&d(\wedge)(x,y)=T \text{ iff } x=T \text{ and } y=T \\
&d(\vee)(x,y)=T \text{ iff } x=T \text{ or } y=F \\
&d(\rightarrow)(x,y)=T \text{ iff } x=F \text{ or } y=T \\
&d(\forall)(f)=T \text{ iff } f(x)=T \text{ for every } x \in U \\
&d(\exists)(f)=T \text{ iff } f(x)=T \text{ for at least one } x \in U
\end{aligned}
$$

Remaining constraints can be articulated as follows:

$$
\begin{aligned}
&d(\mathbb{R}_{O1}(o,S)) = d(o)(d(S)) \\
&d(\mathbb{R}_{O2}(o,S_1,S_2)) = d(o)(d(S_1),d(S_2)) \\
&d(\mathbb{R}_{Pn}(P^n,T,\dots,T)) = d(P^n)(d(T),\dots,d(T)) \\
&d(\mathbb{R}_Q(Q,P)) = d(Q)(d(P)) \\
&d(\mathbb{R}_\lambda(S,T)) \text{ is such a function f that } f(d(T')) = d(S^{[T' \rightarrow T]}) \text{ for every term } T'
\end{aligned}
$$

(Only the first four of these constraints really are of type (iii) - what the last one says is clearly *not* how the denotation of the value of $\mathbb{R}_\lambda$ depends on the denotations of its arguments. To be able to make do only with the constraints of type (iii) we would need a much more complicated grammar - we would have to replace $\mathbb{R}_\lambda$ by a family of rules such as those of Quine, 1960b. But this is not important here and we are not going to do it.)

The last point to conclude the definition of denotation-assignment is the definition of a subset of the domain of denotations of statements, i.e of those facts that are to count as actual. This is quite straightforward - the truth value *truth* counts as the single actual fact. This means that if S is a statement, then $d(S) \epsilon A$ if and only if $d(S)=T$.

7.4 Compositionality and Domains

What is the significance of this for the general case? How could one arrive at constraints of the types (i)-(iii) in the case of a more general logical calculus?

Let us start with constraints of type (i). What kind of entities should in general be assigned to expressions of various categories? At first sight there seems to be no clear general reason (if we leave aside the speculations mistaking model theory for metaphysics) why expressions of some grammatical categories should be taken as denoting objects of a special kind, for example some kind of functions. But for the traditional concept of denotation of first-order logic, just this seems to be essential: expressions of all categories, with the exception of terms and statements, are interpreted by functions.

Why are, e.g., predicates of the predicate calculus taken to denote functions from individuals to truth values (i.e. classes of individuals)? The answer of a formal metaphysician might be: because predicates *really do* denote such functions. However, there is a simpler, non-metaphysical answer: the reason is that if we take predicates to denote such functions, then we facilitate easy computability of the denotations of atomic statements (i.e. of their truth values) out of the denotations of their parts.

An atomic statement consists of an n-ary predicate and n terms; hence the denotation of an n-ary predicate plus those of n terms must "give" the denotation of a statement. This is to say that we must assume the existence of a functor which maps every denotation of an n-ary predicate plus the denotations of n terms on the denotation of the statement which results from the combination of the predicate with the terms. In other words, we have the grammatical rule $\mathbb{R}_{Pn}$, and hence we must assume the existence of its denotational companion $\mathbb{R}_{Pn}^*$ such that $d(\mathbb{R}_{Pn}(P,T_1,\ldots,T_n)) = d(P(T_1,\ldots,T_n)) = \mathbb{R}_{Pn}^*(d(P),d(T_1),\ldots,d(T_n))$ for every n-ary predicate P and terms $T_1,\ldots,T_n$. If we consider statements as denoting truth values, then $\mathbb{R}_{Pn}^*$ can be considered as an (n+1)-ary functor mapping every (n+1)-tuple consisting of a denotation of an n-ary predicate plus the denotations of n terms on a truth value. The existence of $\mathbb{R}_{P1}^*$ is all we are really bound to assume; there is no principal reason to identify denotations of predicates with functions from denotations of terms to denotations of statements. If, however, we do so, i.e. if we take $d(P)$ to be the function such that $\mathbb{R}_{P1}^*(d(P),d(T_1),\ldots,d(T_n)) = d(P)(d(T_1),\ldots,d(T_n))$, then we in a sense push $\mathbb{R}_{P1}^*$ "inside" the denotations of predicates; and everything is considerably simplified. We then always have $d(P(T_1,\ldots,T_n)) = d(P)(d(T_1),\ldots,d(T_n))$.

The situation is similar for other grammatical rules. Let us consider $\mathbb{R}_{O1}$: all we are bound to assume is the existence of a $\mathbb{R}_{O1}^*$ such that $d(\mathbb{R}_{O1}(o,S)) = \mathbb{R}_{O1}^*(d(o),d(S))$ for every unary operator o and every statement S; however, if we again assume that $d(o)$ is the function such that $\mathbb{R}_{O1}^*(d(o),d(S)) = d(o)(d(S))$, everything is much simplified. This indicates that constraints of the type (i) must be considered as inseparable from those of type (iii): the purpose of the former can in fact be seen in providing for the simplification of the latter.

This means that constraints of the types (i) and (iii) must be seen as interlocked - constraints of type (i) provide for the simplicity of the functors stipulated by constraints of type (iii). Constraints of type (iii) guarantee that the denotation of a complex expression will always be uniquely determined by the denotations of its parts; and constraints of type (i) aim at making the way in which the denotation of the whole depends on those of its parts as perspicuous as possible. We can make the companions of all but one of our grammatical rules into such simple "applicators"[88] in that we make denotations of certain categories into functions - only the companion of $\mathbb{R}_\lambda$ remains more complicated. This brings about what can be called the *functional paradigm* of traditional logic - denotations of expressions of all categories save some basic ones (statements and terms) are considered as functions, and the rules of composition of

[88]Where an n-ary functor F can be called an *applicator* if and only if there is an i such that $1 \leq i \leq n$ and if $x_1,\ldots,x_n$ are from the domain of F, then x_i is a function such that $F(x_1,\ldots,x_n)=x_i(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)$.

denotations are taken to be the rules of function application. This functional paradigm finds its most general expression in the language of the simple theory of types as articulated by Church (1940) and Kemeny (1948).

What about constraints of the type (ii)? It is clear that they can be considered as resulting from the corresponding constraints on satisfaction (while the constraints on satisfaction can in turn be seen as resulting from an articulation of inferential patterns - see Section 4.4). In Section 7.1 we formulated the constraints ($C\wedge''$), ($C\vee''$) and ($C\rightarrow''$) - now, using the facts that $\mathbb{R}_{O2}{}^{*}(d(o),d(S_1),d(S_2)) = d(o)(d(S_1),d(S_2))$ and that - in the case of extensional logic - $d(S)\epsilon A$ if and only if $d(S)=T$, these constraints can be turned into ($C\wedge'''$), ($C\vee'''$) and ($C\rightarrow'''$), and thus in fact into the constraints on $\wedge$, $\vee$ and $\rightarrow$ which are listed in the previous section; the same holds for constraints concerning other logical constants.

$$d(\wedge)(d(S_1),d(S_2))=T \text{ iff } d(S_1)=T \text{ and } d(S_2)=T \qquad (C\wedge''')$$
$$d(\vee)(d(S_1),d(S_2))=T \text{ iff } d(S_1)=T \text{ or } d(S_2)=T \qquad (C\vee''')$$
$$d(\rightarrow)(d(S_1),d(S_2))=T \text{ iff } d(S_2)=T \text{ or } d(S_1)=F \qquad (C\rightarrow''')$$

7.5 Modal Contexts and Possible Worlds

The language of modal logic can be considered as arising from that of classical logic by adding a single constant, namely $\Box$, regimenting roughly the natural language expression *necessarily*. However, the inference rules brought about by the introduction of the constant necessitate deep changes in the resulting model theory: it is no longer satisfactory to consider statements as denoting truth values, we need a richer repertoire of proto-facts.

This threatens the whole functional paradigm envisaged by classical logic. The situation can, however, be saved, at least to a large extent. How to do so has been shown by Kripke (1963). It is enough, Kripke demonstrated, to consider statements as denoting subclasses of a class, the elements of which he called *possible worlds*; and to adjust denotations of expressions of other categories accordingly.[89] The resulting model theory is still reasonably elegant and reasonably "functional" (in the sense of the functional paradigm).

It is important to realize that there is an analogy between positing possible worlds and positing individuals as discussed in Section 5.7.[90] It is clear that accepting Kripkean semantics, we can, and we usually do, construe $\Box$ as implicitly quantifying over possible worlds. So similarly to our taking the falsity of $\forall xS$ as the proof of the existence of an individual counter-satisfying S, we can now take the falsity of $\Box S$ to prove that there is a possible world w.r.t. which S is false. We can consider classical logic as restricted to a single world (namely the actual one); and we can treat the fact that $\Box S$ can be false despite S being true as proving that there are possible worlds beyond the actual one.[91] By multiplying worlds we multiply also proto-facts (which can be - in accordance with Kripke's proposal - construed as classes of possible worlds); so whereas classical logic worked with only two proto-facts (the truth value *truth*, or "The Great Fact" and the truth value *falsity* or "The Great Fiction"), the introduction of $\Box$ means the essential refinement of the domain of proto-facts.

In Chapter 5 we saw that in some particular cases we can "dis-necessitate", i.e. we can reduce the necessary truth of a statement to simple truth of another - quantificational - statement. $\Box$ provides for "dis-necessitation" in general: the necessary truth of S is always - by definition - the same as the simple truth of $\Box S$. This means that instead of studying necessary truth, which requires considering the whole space of admissible truth-valuations,

---

[89]In Section 4.5 we considered possible worlds as model structures; so this is a second notion of possible world. Unlike model structures, Kripkean possible worlds are initially, as Stalnaker (1986, p.117) puts it, "not a particular kind of thing or place" but rather "what truth is relative to, what it is the point of rational activities such as deliberation, communication and inquiry to distinguish between." However, this is not to say that the two notions would be incompatible - insofar as truth in every model structure (the validity of S) coincides with truth w.r.t. every possible world (the truth of $\Box S$), the two notions naturally coalesce. Cf. Peregrin (1993a).

[90]The analogy is quite obvious if we look at the standard way of generalizing of Henkin's (1950) completeness proof to intensional logic (e.g., Gallin, 1975, p.25; lemma 3.2).

[91]See Peregrin (forthcoming d).

we can restrict ourselves to considering truth *simpliciter*, i.e. only the actual truth-valuation.

The inclusion of modal aspects of natural language into the scope of logical schematization thus has far-reaching consequences. First, it makes denotations fine-grained in such a way that the relation of sameness of denotations is no longer entirely remote from the intuitive relation of synonymy.[92] Second, it makes denotations more like meanings also in another respect - it is now possible to restrict ourselves to a single denotation-assignment instead of having to consider a whole space of them.

7.6 The General Character of Denotation

Let us now try to articulate those general principles which have so far been found to underlie the concept of denotation and which are manifested within usual model theories as developed for various extensions and modifications of traditional logic[93]. There seem to be several general principles whose embodiments the particular denotation-assignments as defined for the particular logical calculi appear to be.

First, denotation-assignment always appears to be compositional. This means that the denotation of every complex expression is uniquely determined by the denotations of its parts; i.e. that the denotation of the output expression of a grammatical rule is always a function of the denotations of its input expressions. This can be called *the principle of compositionality*. This principle is clearly equivalent to what can be called *the principle of intersubstitutivity*:

[92]Let us note that the refinement brought about by the traditional way of introducing possible worlds is usually considered as not generally satisfactory - the resulting proto-facts are not considered sufficiently fine-grained to allow for the adequate treatment of reports of propositional attitudes; and the general trend is to abandon the functional paradigm of traditional logic (see Peregrin, 1993a). However, let us note that there is also a way of accounting for the propositional attitudes not eschewing the possible world framework - see Hintikka (1978).

[93]Modal logic (Kripke, 1963; Lemmon, 1966); higher-order predicate logic up to Church's (1940) calculus of the simple type theory (Kemeny, 1948; Henkin, 1950); intensional logic (Montague, 1974; Gallin, 1975) etc.

replacing part of a complex expression by another part with the same denotation cannot change the denotation of the whole expression.

Second, the proto-facts (i.e. the entities - whatever they might be - denoted by statements) always, at least in principle, fall into two categories, into "actual" and "unactual". This means that statements differing in truth value are bound to differ in denotation - true statements denote actual facts, false statements the unactual ones. So there appears to be a principle claiming that denotation is grounded in truth; let us call it *the principle of verifoundation*.

Compositionality together with verifoundation entail *the principle of salva-veritate-intersubstitutivity*: replacing part of a statement by another part with the same denotation cannot change the truth value of the whole expression. Moreover, if we extend the concept of part in such a way that every expression is also considered its own part, then the inverse holds too: *salva-veritate*-intersubstitutivity implies compositionality and verifoundation. (However, note that this would not necessarily be true if we understood compositionality in a constructivistic sense - as claiming that the denotational companions of grammatical rules are some kinds of effective procedures.)

The two principles just stated render the core of the notion of denotation; however, if they were the only two principles governing the notion, then the function mapping every expression on itself could count as denotation-assignment - and this seems to be odd. The characteristic feature of denotation seems to be that it passes over all differences save the indispensable ones - nominally different expressions do not necessarily differ in denotation, they differ only when there is a "semantical reason" for them to differ. This amounts to *the principle of 'Occam's razor'*: two expression differ in denotation only when the opposite would contradict compositionality or verifoundation. Thus, denotation renders the "most economic" compositional and verifounded assignment of objects to expressions in the sense that any other compositional and verifounded assignment differentiates at least as much as denotation.

We have seen that compositionality plus verifoundation are equivalent to *salva-veritate*-intersubstitutivity; Occam's razor is now equivalent to a principle inverse to *salva-veritate*-intersubstitutivity: if two expressions are intersubstitutive *salva veritate*, then they have the same denotation. This means that compositionality, verifoundation and Occam's razor are together equivalent to *the Leibniz principle*: two expressions (of the same grammatical category) have the same denotation if and only if they are intersubstitutive *salva veritate*. Hence the principles determine denotation uniquely "up to isomorphism"; they uniquely determine the relation of sameness of denotation, but leave open the character of the objects being denoted.

So there is still a degree of freedom; and this allows for the formulation of a fourth principle, which is more pragmatic, and also more vague, than the previous three. We can call it *the principle of simplicity of composition*: the ways in which the denotation of a complex expression is obtained from those of its parts is "as simple as possible". In some cases we can turn this principle into a more specific requirement: denotation is "applicatively" compositional. The denotation of a complex expression is the value of the denotation of one of its parts applied to those of the other parts.

These four principles provide for a non-metaphysical way of explaining of why model theory of the common logical calculi looks the way it does - it is not that individuals, relations, etc. were the best renderings of the way the world behind language is, it is rather that items which have come to be called this way have proven themselves to be the best tools for gripping onto truth and consequence.

If we return to our algebraic watchtower, we can perceive the principles as follows. The principle of compositionality says that denotation-assignment is a homomorphism of the language-algebra into the algebra of denotations. The principle of verifoundation claims that the kernel of a denotation-assignment (i.e. the relation connecting those pairs of expressions whose denotation is the same) is a congruence w.r.t. the relation of sameness of truth values among statements. The principle of Occam's razor then stipulates that such a congruent homomorphism counts as a denotation-assignment only if its kernel is maximal, in the sense that it is not contained in that of another congruent homomorphism. It is easily seen that a homomorphism is a denotation-assignment if and only if its kernel coincides with the relation of intersubstitutivity *salva veritate*. The principle of simplicity then states that the operators of the algebra of denotation are in some sense maximally simple.

8 Meaning

8.1 The Nature of Meaning

The concept of meaning is grounded in our pre-theoretical feeling that to take a string of letters as an *expression* is to take it as standing for something. A proper name seems to be the paradigmatic example: it gets explicitly attached to a definite individual during the act of christening, its purpose thereafter being to point out the person to which it is attached. Other linguistic expressions then are imagined as getting attached to other kinds of objects - maybe to objects more obscure than people (e.g. ideas, facts, functions, representations, or whatever), and maybe in ways less explicit and unequivocal than christening (e.g. by some obscure stepwise processes taking place during the forgotten prehistory of mankind) - but in principale in the same manner. As Tichý (1992) puts it, language is a code and a theory of language thus has to consist in cracking the code.

In fact, the concept of meaning needs to do justice to *two* basic intuitions. The meaning of an expression seems to be (i) what the expression intuitively "stands for"; and (ii) what makes the expression "expressive", i.e. what makes it usable for the purposes of communication. The code-conception of language suggests that these two requirements are two sides of the same coin - expressions are nothing over and above substitutes for objects of the real world, and hence that which makes an expression capable of serving the purpose of conveying messages is simply that it stands for an object.

The plausibility of such a picture derives from the fact that many words really appear to be used as names of the objects we are familiar with from everyday life. Thus the name *Fido* stands for my dog Fido, the description *the dog I found last December forsaken in the railway station* for the very same dog, and the word *dog* stands ambiguously for any dog, or, using the achievements of formal logic, unambiguously for the set of all dogs. However, as was pointed out by Frege and many other theoreticians, if we accept (ii), then, in the majority of cases, the meaning of an expression simply cannot be the object the expression is felt to stand for. If both *Fido* and *the dog I found* *last December forsaken in the railway station* meant the dog Fido, then we would be free to use the two expressions interchangeably. (The point is that (ii) implies that expressions coinciding in meaning are wholly indifferent from the point of view of communication; and given (i) this means that expressions standing for the same object can never differ in their communicative function.) Thus, to utter *Fido is the dog I found last December forsaken in the railway station* would be the same as to utter *Fido is Fido*.

Frege (1892) attempted to resolve the conflict by splitting the concept of meaning into two parts: into the concept of reference (*Bedeutung*) which does justice to the pre-theoretical intuition concerning "standing for", and the concept of sense (*Sinn*) that saves the situation when reference alone is incapable of accounting for all those properties of the expression which appear relevant from the point of view of communication. Thus, *Fido* and *the dog I found last December forsaken in the railway station* share the common referent, my dog Fido, but they differ in sense, in the way they present the referent. The nontrivial communicative function of *Fido is the dog I found last December forsaken in the railway station* lies in its telling us that the two senses expressed by the two noun phrases point out the same referent. So we can understand Fregean sense as capturing all that which there is to meaning beyond reference.

However, the possible discrepancy between (i) and (ii) may be deeper than can be dispensed with by Frege's manoeuvre (not to speak of the fact that Frege's proposal is really plausible only for certain expressions - names and, possibly, statements - not for expressions of other grammatical categories). The point is that (i) suggests that the relationship between expression and its meaning is a matter of a casual attachment of two otherwise independent entities; whereas (ii) rather suggests that meaning is a kind of hypostatic capturing of the function of expression within the process of communication. The approach to meaning based on taking (i) as crucial takes semantics to be a species of semiotics, of the science dealing with ways of letting objects stand for other objects; the approach stressing the essentiality of (ii), on the other hand, considers semantics as a matter of the study of the structure of communication and of language.

8.2 Two Approaches to a Theory of Meaning

So we have two essentially different kinds of approaches to meaning. One is based on the assumption that expressions are more or less conventional substitutes for objects of the real world; that the significance or meaning of an expression is a matter quite independent of the fact as to whether the expression is a part of language or of some other system of signs. The second approach, on the other hand, is based on the assumption that any significance, or meaning, that an expression acquires is the result of the expression's being part of the system of language (i.e. of its systematic involvement with the praxis of communication).

The first of these approaches thus tries to explain meaning in terms of the relation of *representation* - the meaning of an expression is what the expression represents; hence something is an expression inasmuch as it represents, as it acts as a substitute for something else, for something existing independently of the expression. Viewed in this way, the relationship between an expression and its meaning seems to be explicable on the causal basis - the expression means something because it was somehow caused to stand for it. Language is thus seen as a nomenclature of objects; its structure, especially the structure created by the relation of consequence, is only some kind of reflection of the structure of the world it represents.

The second view stems from the reluctance to see an expression as a sign of its meaning in the sense in which a label on a drawer in a drugstore is the sign of what the drawer contains; or in which the red light is the sign for stopping. According to this view, language is not a nomenclature of real-world objects and to view it in this way mean to fall with what Quine (1969) calls the *museum myth*. Whatever it is that is signified, the signification relation is derived from the relevant interrelations of expressions within the system of language. The meaning of an expression is yielded by its involvement with the communicative praxis, especially with the praxis of making inferences. Thus, it is the structure of language which is primary, and the relation between an expression and its meaning is what is derived.

If we accept the picture of language proposed long ago by de Saussure (1931), we can see the relations linking expressions to other expressions (thus amounting to the structure of language) as *horizontal*, while those linking expressions to their meanings as *vertical* (see Figure 2) and we can say that the two approaches differ in which of these relations they grant primacy. The first of the approaches maintains that it is the vertical relations that are primary; that any relations between expressions are derivative to the relations of the expressions to their meanings. The second approach, on the other hand, grants primacy to the horizontal relations, it maintains that it is the systemic relations between expressions that are primary, and that the relation between an expression and its meaning is what is derived.

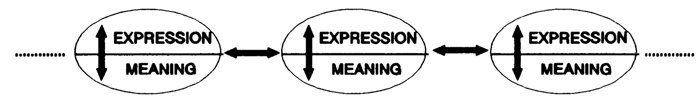

Figure 2

We may call the first approach *nomenclatural* and the second *structural*[94]; or, if we want to stress that the structure of language that is mainly relevant is its inferential structure we may - borrowing from Brandom (1985) - call the first *representational* and the second *inferential*[95]. The nomenclatural or representational view is surely closer to the commonsense idea of meaning; and it is also accepted by the mainstream of philosophers. It is, however, becoming more and more clear that this view is untenable. Within the tradition of analytic philosophy, the first clear evidence against the view follows from Frege's recognition of the fact pointed out in the previous section, namely that one and the same entity taken as meaning cannot in general do

[94]Cf. Peregrin (1993b).

[95]Brandom (1985) gives an insightful discussion of the history and background of the conflict of the two approaches. His claim is that "the philosophical tradition can be portrayed as providing two different models for the significances which are proximal objects of explicit understanding, representational and inferential. We may call 'representationalism' the semantically reductive view that inference is to be explained away in favor of more primitive representational relations. ... By 'inferentialism', on the other hand, one would mean the complementary semantically reductive order of explanation which would define representational features of subsentential expressions in terms of the inferential relations of sentences containing them." (p.31)

justice to both the intuitions (i) and (ii). Another fact questioning the nomenclatural view, a fact which was also noticed by Frege and which was then pointed out by Wittgenstein, is that the elements of language which are primarily significant are not words, but rather sentences[96]. It follows that the discernment of the senses of parts of a sentence is "parasitic"[97] upon the structure of the sentence and that it is, therefore, the structure of language that is constitutive to individual senses.

The nomenclatural view of meaning is thus essentially implausible. It may, however, seem that it can be defended at least with respect to reference. However, Quine (1960a) has pointed out that if reference amounts to the commonsense idea of "standing for", then the very possibility of our determining it definitely is precluded; and Davidson (1977) took the ultimate step of showing that such a concept of reference - the concept of reference that is considered to do justice to the intuition (i) - is completely meaningless.

8.3 Expressions and Chess Pieces

Probably the most illuminating metaphor used to elucidate the notion of meaning resulting from the rejection of the nomenclatural notion is the metaphor comparing the meaning of an expression with the role of a chess piece. The affinity between chess and language is due to the fact that the significance of their elements (the pieces or the expressions) does not consist in their peculiar nature (in their shape or in the material they are made of), but rather in the role they play in the system. De Saussure (1931; p.88,110), whose name is usually associated with structuralism, puts it in the following way:

> A game of chess is like an artificial realization of what language offers in a natural form. ... Take a knight, for instance. By itself

---

[96] "Nur im Zusammenhange eines Satzes bedeuten die Wörter etwas." ["It is only in the context of a proposition that words have any meaning."] (Frege, 1884, p.73). "Nur der Satz hat Sinn; nur im Zusammenhange des Satzes hat ein Name Bedeutung." ["Only propositions have sense; only in the nexus of a proposition does a name have meaning."] (Wittgenstein, 1922, §3.3).

[97] In Dummett's (1988) term.

> is it an element in the game? Certainly not, for by its material makeup - outside its square and the other conditions of the game - it means nothing to the player; it becomes a real, concrete element only when endowed with value and wedded to it.

However, this view of language is far from restricted to de Saussure and his avowed structuralist followers. In fact it is characteristic of what we have called the structural standpoint, and cuts across usually acknowledged philosophical schools, making de Saussure closer to Wittgenstein than to those who are usually considered as "structuralists"[98]. No wonder then, that Wittgenstein (1953, §108) describes language in almost the same terms[99]

> Wir reden von dem räumlichen und zeitlichen Phänomen der Sprache; nicht von einem unräumlichen und unzeitlichen Unding. Aber wir reden von ihr so, wie von den Figuren des Schachspiels, indem wir Spielregeln für sie angeben, nicht ihre physikalischen Eigenschaften beschreiben. Die Frage "Was ist eigentlich ein Wort?" ist analog der "Was ist eine Schachfigur?"[100]

Wittgenstein's view of language is based on the assumption that what makes a sequence of letters an expression is not a kind of causal connection with an extralinguistic object, but rather its incorporation into the system of language. From this point of view, Wittgenstein is a structuralist and an inferentialist[101]; and so are some of the more recent analytic philosophers,

---

[98]See Peregrin (1995).

[99]Note too that chess metaphors were also employed by Frege (1885) and Husserl (1900/1); not to mention their followers.

[100]We are talking about the spatial and temporal phenomenon of language, not about some non-spatial, non-temporal phantasm. But we talk about it as we do about pieces in chess when we are stating the rules of the game, not describing their physical properties. The question "What is the word really?" is analogous to "What is a piece in chess?"

[101]The Wittgensteinian inferentialist's standpoint can be elucidated by the observation made by Malcolm (1940, p.197): "Philosophers and logicians have the idea that when a question as to whether one statement entails another arises, verbal considerations enter only because of ambiguity, and that the *real* question is not a verbal one, but one to be settled by the intellect's fixing its gaze upon the proposition, *after* the ambiguity has been cleared up

notably Quine and Davidson.[102] In a recent writing, Quine (1992, p.9) makes his structuralist *credo* quite explicit: "Save the structure and you save all."[103] However, the structural insight is present in Quine's writings from the beginning - in fact it underlies the most essential part of his teaching, namely his system of interlocked indeterminacies[104]. Let us illustrate the problem by a simple example.

Let us imagine that a speaker S_1 uses the word 'N' to refer to an object A, while his fellow speaker S_2 misuses the same name to refer to another object B. How can S_1 find out about S_2's misuse? If A and B can be unequivocally pointed at, then there is, of course, no problem; but this is usually not the case. However, even if there is no direct pointing at, the resolution seems to be simple; for S_1 can simply say: "What you call 'N' is B, but 'N' is in fact the name of A!" This possibility is nevertheless limited to the case when both speakers are in possession not only of the name 'N', but also of the names 'A' and 'B'. However, in this case, 'N' is clearly superfluous and the whole situation is uninteresting. What about the case when they have no such names?

In such a case, S_1 cannot, of course, say "What you call 'N' is not N!"; because for S_2 what he calls 'N' surely *is* N. However, if there is something that is true for A, but not for B, then S_1 can make use of this fact. If, e.g., A is green, whereas B is not, then S_1 can put forward the statement "N is green"; S_2 will disagree, and S_1 and S_2 then have to conclude that they use 'N' in different ways (presupposing that neither of them is colour-blind).

Thus S_2's misuse of 'N' may be discovered as soon as the speakers encounter a statement containing 'N' and such that it would be true in case 'N' were A and false in case 'N' were B. If there is no such statement (and if A

... [However,] clearing up the ambiguity of a statement *consists* in showing what it means, and this *consists* in showing how it is used, and this *consists* in showing what it entails."

[102]The scholars of the opposite wing of the analytic movement embrace representationalism: they take the relation of expression's standing for an object as primary - they take it either simply as a causal relation (drawing on the analyses of Kripke, 1972, Putnam, 1975, and others), or as a relation composed out of a simple causal relation of the expression to a content of one's consciousness, and an irreducible relation of the content of consciousness to what the content is about (Schiffer, 1972, Searle, 1983, Fodor, 1987).

[103]See also Peregrin (forthcoming a).

[104]Indeterminacy of translation, inscrutability of reference, ontological relativity - see Quine (1969).

and B cannot be distinguished by direct ostension), S_1 can never find out that S_2 is talking about another object. This is the case of Quine's (1960a) celebrated example of a rabbit, its undetached parts and its stages.[105]

What does this example show? One possibility is to conclude that in some cases it is essentially impossible to find out what one is talking about; but this is an interpretation which is not available for a theoretician who views language in the structural way. For him, any "aboutness" of an expression is grounded in the way the expression functions in language. So the structuralist is bound to conclude that in many cases it simply does not make sense to consider a speaker's "speaking about" or a word's "standing for" as a determinate relation between two determinate objects.

However, did not our example lead us to the clear conclusion that S_1 and S_2, when using 'N', simply cannot know what they are talking about - a conclusion we assume to be unacceptable for the structuralist? Well, the structuralist's response is that the absurd conclusion is in fact the *reductio ad absurdum* of the premise, namely that S_1 and S_2 are using 'N' to refer to different objects. If S_1 and S_2 use 'N' in the same way, then it is absurd to say that they use it to refer to different things. If the thing referred to by an expression is irrelevant for the way language functions, then there is no sense in considering reference and things referred to as a concern for a theory of language. Davidson (1977, p.225) concludes: "Reference ... drops out. It plays no essential role in explaining the relation between language and reality."[106]

Moreover, from the point of view of S_1 and S_2 there are no such two entities as A and B at all; there is a single entity N. To talk about A and B as about two different things makes sense only because we have a theory that renders them different. It is only from *our* viewpoint that it makes sense to say that S_1 and S_2 use 'N' ambiguously between A and B. Individuation is a matter of (such or another) theory. To posit a difference makes sense only on a background of a language which is capable of articulating it. Thus, Quine (1969, p.51) concludes that "...talk of subordinate theories and their ontologies

---

[105]Note also that Quine's example indicates that unequivocal ostension is impossible even in the case of the most ordinary physical things.

[106]"Das Ding in der Schachtel," we can say with Wittgenstein (1953, §293), "gehört überhaupt nicht zum Sprachspiel; auch nicht einmal als ein *Etwas*: denn die Schachtel könnte auch leer sein." ["The thing in the box has no place in the language-game at all; not even as a *something*; for the box might even be empty."]

is meaningful, but only relative to the background theory with its own primitively adopted and ultimately inscrutable ontology." And Davidson (1979, pp. 233-234) adds that "if this is the situation, ontology is not merely 'ultimately inscrutable', any claim about reference, however many times relativized, will be as meaningless as 'Socrates is taller than'."[107]

This is also precisely what contrasts the structural with the nomenclatural view. If we accepted the nomenclatural view, then we could not see any problem in distinguishing whether 'N' stands for A or B even if no evidence at all stemmed from the way 'N' functions within language; it would simply refer to that object to which it has once been causally attached. Not so within the structural view: we may differentiate vertical links only in terms of horizontal ones, and where there is no difference on the horizontal level, there cannot be a difference on the vertical one. Any difference between meanings of two expressions must consist in a difference between the functioning of the expressions in language; to posit any difference over and above this simply does not make sense.

However, let us once more stress that it would be essentially wrong to interpret this structuralist inscrutability as implying that we cannot know what our expressions mean or what they "stand for"; rather the right interpretation is that meaning and reference are not definite in the sense presupposed by the nomenclatural view. There is, for example, little doubt that natural numbers exist; at least as much as objects which cannot be touched and tasted can exist. We consider them as unique existing objects although we know that for all we care they can be identified with objects of various kinds[108]; a number is not one of the things it can be reduced to, it is rather what all of these things have

---

[107]Putnam's *internal realism*, which is in fact nothing else than a way to cope with *ontological relativity*, brings about a similar diagnosis: "What looked like an innocent formulation of the problem - 'Here are the objects to be referred to. Here are the speakers using words. How can we describe the relation between the speakers and the objects?' - becomes far from innocent when what is wanted is not a 'natural-language processor' that works in some restricted context, but a 'theory of reference'. From an internal realist point of view, the very problem is nonsensical." (Putnam, 1988, p.120)

[108]Certain classes of classes of real-world objects, or certain classes built recursively out of the empty class; to mention at least the most popular proposals.

in common, it is *a number*[109]. And what can be said about numbers, can, more generally, be said about meanings (numbers are, after all, meanings of numerals and numerical expressions): meanings are not definite things conforming to our usage of language, we can at most say that they are what all such things have in common. As de Saussure (1931, p.122) puts it, "language is a form and not a substance".

8.4 Formal Semantics vs. Formal Metaphysics

The formal explication of the concept of meaning is now usually considered as a mapping of language expressions on set-theoretically modelled extralinguistic objects. This may, and it often does, make one conclude that to do formal semantics we must first set-theoretically reconstruct the world, and only then can we pair expressions with appropriate objects resulting from the reconstruction; i.e. that prior to doing formal semantics we must do a kind of formal metaphysics. This conclusion is a necessary result of the nomenclatural view of language - we must first expound and catalogue the things which are labelled by expressions, the exhibits of the world-museum, and only then can we inquire into which label is stuck on which thing.

If we view language structurally, we cannot see any point in basing the theory of meaning on a metaphysics. The way of metaphysics, including that of formal metaphysics, is doomed, because of the lack of intersubjectively shared criteria - as was clearly shown also by pioneers of analytical philosophy. Moreover, viewing the meaning of an expression as a kind of reification of the way the expression functions within language allows us to see semantics as based on the reconstruction of language, not in a reconstruction of the language-independent world. Thus, instead of considering the theory of language as supplementary to the theory of the world, instead of considering the structure of language as derived from the structure of the world, the structuralist sees the situation the other way round; he wants to analyze language and then possibly use the results to elucidate the world.

[109]As Quine (1969, p.45) puts it: "there is no saying absolutely what the numbers are; there is only arithmetic".

However, let us note that by rejecting the nomenclatural view of language and by rejecting metaphysics as a descriptive theory of the world independent of language, we do not claim that there can be nothing that could be called formal metaphysics. The model-theoretical way of accounting for necessary truth, as exposed above, may lead us to positing ("auxiliary") entities whose natural names suspiciously remind us of those of entities of traditional metaphysics: individuals, properties, facts; and maybe also events, situations, possible worlds. The presentation of the inferential structure of language using such entities may then be naturally considered as doing a kind of metaphysics; and we may begin to speak, together with Bach (1986), about *natural language metaphysics*. However, this sense of doing metaphysics does not amount to cataloguing the language-independent world, but rather to expounding the structure of language; nevertheless if we take the results of the philosophers of the linguistic turn seriously, we have to conclude that this is indeed the only kind of metaphysics that can be really done.[110]

The difference between the two approaches to formal semantics can be illustrated using two different 'metaphysical theories': namely Davidson's (1967b) theory of events and Barwise & Perry's (1981) theory of situations, respectively. Davidson starts from the analysis of action sentences and takes pains to show that we can clearly and comprehensibly account for the inferential behaviour of these sentences if we analyze action verbs as predicates containing a covert argument place which gets existentially quantified on the level of the whole sentence. He proposes to call the items which are so quantified over *events*. The contribution of Davidson's analysis of action sentences to the question *what is there?* could therefore be considered to be the claim *there are events*. However, suppose that someone protested: *there are no events*. Davidson's way to defend his claim would not be an appeal to intuition, to introspective analysis of one's consciousness nor to extrospective analysis of the world, his way of arguing for the existence of events would have to be

[110]Bach's (1985, p.593) confession that the main difference he is able to see between philosophers doing metaphysics and linguists doing model-theoretical semantics "lies in whether you are embarrassed about not knowing about a paper in *Linguistic Inquiry* or the *Journal of Philosophy*" is more than a mere *bonmot*. Bach clearly shows that analyzing language yields us results that strike us as relevant to the question *what is there?* and that a linguist thus cannot escape doing a kind of metaphysics - it is, however, metaphysics resulting from parsing language, not from parsing the world in a direct, non-linguistic way.

based on the fact that events are simply what we require for composing a comprehensible model-theoretical account for a certain part of language.[111]

The approach of Barwise & Perry, although seemingly similar, is in fact essentially different. Their way is to first develop the "Theory of Situations", an "abstract theory of talking about situations" based on "pulling out of the real situations the basic building blocks of the theory", and then to use this theory as the basis for a theory of semantics.[112] Barwise & Perry's answer to the question *what is there?* can therefore be considered as the claim *there are situations* and might thus seem to be of the same kind as Davidson's answer. However, the two answers are crucially different: in contrast to Davidson's, Barwise & Perry's way of rejecting the sceptic claiming *there are no situations* would have to be based on speculating about the structure of the world or of our way of perceiving it.

The Davidsonian way of doing 'metaphysics', unlike that of Barwise & Perry and many of contemporary analytic philosophers and formal semanticians, consists in bringing to light the inferential structure of language. This accords with the view of model theory we adopted above: we see model-theory as the unfolding of certain systemic properties of language, not as an imitation of the way words hooks on things.

8.5 Truth and Meaning

If we account for truth in terms of the Tarskian apparatus of satisfaction, and if we turn, as we did in the previous chapter, satisfaction into denotation, then what we gain is a "reification" of truth in the form of entities, denotations, assigned to expressions. What we propose is that it is just these entities that can be adequately considered as meanings. This implies that our technical concept

---

[111]In his book, elaborating the Davidsonian event-based approach to English in detail, Parsons (1990, p.10) spells out this strategy explicitly: "I begin with a mass of linguistic data to be explained and with the bare outlines of a theory for explaining it. I try to develop the theory in the best way possible to explain the data. Only at the end of the enterprise am I in possession of generalities about events."

[112]See Barwise & Perry (1981, pp.5-8).

of denotation is a good explication of the intuitive concept of meaning. To clarify this let us look more thoroughly at the concept of meaning.

Trying to figure out which rules govern the intuitive concept of meaning leads us to the conclusion that the principles governing the concept of denotation as listed in the previous section are very close to basic intuitive characteristics of meaning. Thus, meaning, as we handle it intuitively, also seems to be a *compositional* matter[113]. Besides this, meaning is what can be called *verifounded*: difference in truth value clearly implies difference in meaning[114]. Third, there seems to be something like *Occam's razor*, something that pushes the number of meanings as far down as possible[115]. This indicates that the pre-theoretical notion of meaning-assignment can be explicated as the theoretical concept of denotation-assignment (in the above sense), as that denotation assignment which is based on the actual distribution of truth values among statements.[116]

The Davidsonian project of analyzing meaning via analyzing truth (roots of which can be traced back to Frege) is the project of singling out the individual contribution an expression brings to the truth value of the statements it is a part of, and of identifying the meaning of the expression precisely with this individuated contribution.[117] Now the principles stated in the preceding section to characterize the notion of denotation are exactly what is needed to accomplish such an identification: *the principle of compositionality* and *the principle of verifoundation* state that the contribution of two expressions is not

[113]Compositionality is also what underlies the most influential theories of meaning since Frege. In fact, once we accept Frege's conviction that what is primarily meaningful are sentences, the principle of compositionality becomes necessary for the very individuation of meanings of parts of sentences. It is, however, important to realize that understood this way compositionality is not a thesis to be verified or falsified, but that it is rather a postulate which is constitutive to semantics. See also Janssen (1983).

[114]Cresswell (1982) considers this the most certain principle of semantics.

[115]"Senses are not to be multiplied beyond necessity", as Grice (1989, p.47) puts it.

[116]Where only the truth values of necessary statements really influence meaning. See also Peregrin (1994).

[117]"I suggest that a theory of truth for a language does, in a minimal but important respect, do what we want, that is, give the meanings of all independently meaningful expressions on the basis of an analysis of their structure." (Davidson, 1970, p.55)

the same unless the two expressions are intersubstitutive *salva veritate*, and *the principle of Occam's razor* states that their contribution is the same whenever they are so intersubstitutive. The conjunction of all these principles thus is what Leibniz long ago proposed as a characteristic of meaning; and it is also what can serve as the *operational criterion of synonymy* in the sense of Sgall et al. (1986).

The theory of meaning is thus only a peculiar form of the theory of necessary truth (or consequence); it is, in Quine's (1986, p.115) term, a *spin-of* of the theory of necessary truth[118]. What remains to be explained is what sense there is in addressing truth via meaning.

This question might be given various answers. One answer may be that the model-theoretic account of truth, which has led us to our concept of meaning, is simply a suitable account. Another, popular, answer is that as Gödel's incompleteness results show, it is only the semantic characterization that is capable of accounting for truth adequately. Or it is possible to answer that language is intuitively understood as being *about something*, and that the model-theoretic approach allows us to explicate this intuition.

We would like to propose another answer; namely that it is in a sense in the nature of truth that we must employ meanings to grasp it. This follows from the character of our language. The point is that our language is an infinite set of expressions; but its infinity is a *potential* infinity, it is grounded in the possibility of composing ever more and more complex expressions out of simpler ones. Language thus can be considered as a "constructional" system based on a finite number of urelements (lexicon) and a finite number of constructional rules (grammar)[119]. To say that something is an expression is to say that either it is in the lexicon, or it is a grammatical construction of expressions.

---

[118]As Davidson (1977, p. 222) puts it: "...words, meanings of words, reference, and satisfaction are posits we need to implement a theory of truth."

[119]It is often objected that there are languages (such as, for example, the predicate calculus), which are based on *infinite* lexicons. However, this is misguided: to say that there always must be a finite lexicon in the foundations of any language is not to say that whatever class of expressions we *decide* to call a lexicon would have to be finite. We can, of course, say that the lexicon of the predicate calculus is based on some infinite sets of the kind of $P_1,P_2,\ldots$; but then we must not forget that this is a metaphorical sense of a *lexicon*: in the absolute sense the lexicon is based on the notoriously finite set of several letters (such as P) and the ten numerals.

But if this is the case, then to say what it is for an expression to have some property is to say when elements of the lexicon have the property, and when the results of all the grammatical rules have it. To say which statements are true is to say which simple statements are true and how the truth of a complex statement is to be deduced from the truth of its parts. The problem is, however, that this is impossible: the truth value of a complex statement need not be uniquely determined by the truth values of its parts. The only way to assign truth values to statements is via an intermediary compositional assignment, and this assignment is denotation. Meaning or denotation is thus a product of the "compositionalization of truth".

The last claim, however, may be misleading: it seems to indicate that there is an uncompositional truth and that meaning is added to it to "compositionalize" it. But if we take human finitude at face value, then "uncompositional" turns out to mean as much as "illusory": and this would indicate that there is no truth without meaning. There is no possibility of characterizing truth without saying what the meanings of words are and how meanings of complex expressions depend on meanings of their parts.

Etchemendy (1988, p.76) argues that Tarski's definition of truth does not yield us a theory illuminating the semantic properties of the object language. In contrast, Davidson (1990, p.296) claims that Tarski's construction makes it evident that "for a language with anything like the expressive power of a natural language, the class of true sentences cannot be characterized without introducing a relation like satisfaction, which connects words (singular terms, predicates) with objects" and he insists that this legitimates Tarski's approach as a real theory of semantics. Clearly we side with Davidson, for it is precisely the reification of patterns of inference into objects which is the subject matter of semantics, and which is in the same time - if we disregard oversimplified examples - necessary to survey truth.

The delusion that it is possible to list truths without recourse to semantics stems from Etchemendy's overemphasizing trivial examples of *finite* languages. He claims that we can do a Tarskian theory of truth for such a language without touching its semantics, and that the situation is analogous for infinite languages. But to call a finite collection of signs a *language* is an unhappy metaphor - and expression acquires its meaning, and this is the point of the structuralist insight, precisely in that there are *infinitely many* contexts into which it can be embedded; and its meaning is, as we shall see in detail in the next chapter, precisely an attempt at the summarization of this infinite. We can well say that a finite "language" indeed posseses no semantics over and above that which is captured by its list-like truth definition - this is assuming that we have not already furnish it with semantics by stipulatively equating its expressions to some of those of a *real* language. We shall enlarge on this in the next chapter.

8.6 Davidson's Way

We have characterized the notion of meaning by way of articulating basic constraints this concept seems to be governed by; and in this way we reached the Davidsonian notion of meaning as based on truth. The way is interesting in that it is a way quite different from Davidson's own route. Let us now briefly reconstruct Davidson's own line of reasoning.

Davidson's (1967b) basic assumption is that a theory of meaning for a language should, for every statement s of the language, yield a thesis of the following form

s means m

What should the nature of m in such a thesis be? It should be the name for the meaning of a statement, and such names, Davidson claims, are formed from statements by the prefixing of *that*. So m is to be *that p* for some statement 'p'; (it is clear that if *that p* is to be the name for the meaning of s, then the most straightforward choice would be to have such p that s is 'p'). Anyway, we have

s (='p') means that p

Now the trouble is with the relation 'means that'. The crucial step of Davidson's argument is to assume that this relation can be explicated and eliminated by means of the tools of standard logic; we should "provide the sentence that replaces s with a proper sentential connective, and supply the description that replaces 's' with its own predicate" (p.23). What results is the Tarskian T-sentence.

s (='p') is T if and only if p

The crucial step in Davidson's considerations is thus the elimination of 'means that'; it is this manoeuvre that does the trick of assimilating the theory of meaning to the theory of truth, and it seems possible to insist that it is simply illegitimate, that the relation 'means that' cannot be expressed by means of standard logic and that it is precisely this which prevents the theory of meaning being reduced to the theory of truth. Davidson in fact claims that 'means that' is not a relation between an expression and a thing and that rather it spells out a condition for the expression to have a certain property - an assumption that may be felt as disputable.

To make the point against the view of meaning rejected by Davidson, let us use an example provided by Horwich (ms.). Horwich proposes, for the sake of argument, to consider a notation #p# to indicate the property shared by expressions which rhyme with 'p'; then we can form claims like *'fog' sounds #dog#*. It would clearly be foolish to understand this claim as a relation *x sounds #y#* applied to the objects 'fog' and dog. Now what Horwich in effect argues is that the *that ...* in *s means that p* is just like *#...#* in *s sounds #p#*: it turns *p* into the indicator of the property shared by all expressions synonymous with *p*. *s means that p* is, according to Horwich, not to be considered as a relation between two objects, but should rather be understood as an ascription of a property to an object.

This returns us to where we started the chapter - to the contradiction between the structural and the nomenclatural notions of meaning. The nomenclaturists maintain that to be meaningful means to be linked to an external object, and that *s means that p* expresses such a link. The structuralists, on the other hand, claim that to be meaningful rather means to be in certain way embedded in a structure, to have - or to be - a certain value within the structure. The statement *s means that p*, thus, according to the structuralists, spells out the equi-valence of two expression. It states that the two expressions share their values or, as Horwich puts it, their meaning properties. From this point of view, to say *s means that p* is, as Sellars (1963a) proposed, only a way of saying *s means the same as 'p'* which is in turn only a way of saying *The role of s within its language is the same as that of 'p' within its one*.

8.7 Beyond Truth and Falsity

De Saussure identified language as primarily a system of expressions furnished with oppositions, with entities such as meanings issuing from the system. We have shown how the opposition between truth and falsity (i.e. the opposition between *what there is* and *what there is not*) reifies itself into reference, and how, combined with the opposition between necessity and contingency (i.e. between *what is bound to be* and *what merely happens to be*) it reifies itself into meaning understood as denotation. However, the investigation into the real nature of language reminds us that these oppositions between truth and falsity, although essential, are not the only oppositions which may be relevant for language and for the problem of meaning; and our account in this respect is thus simplified.

The oppositions present in our system of language are related to the ways we use language. The opposition between truth and falsity can be seen as connected with the activity of asserting: true statements are those which can be correctly asserted.[120] But we use language not only to assert something; we may use it to order, to beg or even to accomplish a work of art.

Frege, the early Wittgenstein and a great many of the contemporary analytical philosophers restrict themselves to the study of the assertive capabilities of language. This is reasonable: it is precisely this function of language that is constitutive to our world as the world of things. However, if we want to investigate into the general nature of meaning, then we must remember that making assertions is only one of numerous ways in which we use language.

---

[120]This view may lead us to abandoning the stance which we have maintained so far and which means taking *truth* as a primitive concept. We may, indeed, investigate into the ways we use language and into the norms we establish to govern our usage and show how the concept of truth can be seen as arising out of our practices of asserting and inferencing, in the way pursued e.g. by Dummett (1976) or by Brandom (1983). Such investigations can, and do, bring interesting results. However, I do not think that the possibility to make this switch implies that to treat *truth* as a primitive is an error - for I do not believe that there are any *absolute* primitives. There are only bases we choose, and there is no single concept that could not be excluded from our base. See also footnote 159.

Wittgenstein's later writings make this fact wholly plausible: the "assertive game" is only one of the numerous language-games which can be (and are) "played". Thus while in *Tractatus* we can read[121]:

> Der Satz ist ein Bild der Wirklichkeit. ... Nur dadurch kann der Satz wahr oder falsch sein, indem er ein Bild der Wirklichkeit ist.[122]

in *Philosophical Investigations* Wittgenstein puts forward quite a different view of language and truth[123]

> Und was ein Satz ist, ist in *einem* Sinne bestimmt durch die Regeln des Satzbaus (der deutschen Sprache, z.B.), in einem anderen Sinne durch den Gebrauch des Zeichens im Sprachspiel. Und der Gebrauch der Wörter "wahr" und "falsch" kann auch ein Bestandteil dieses Spiels sein; ...[124]

What Wittgenstein concludes is that if we really want to take all of this into account, we can hardly maintain that there is really something like meaning.

If we, nevertheless, insist on speaking about meaning in this sense, then it must be not only a reification of the expression's usability in the assertive game, but its more general usability for the purposes of the various language games. Meaning of an expression is in general its usability for the various purposes language can serve. To return to de Saussure's terminology, it is *the value of the expression* in the original sense of *value*.

---

[121]Wittgenstein (1922, §4.01, §4.06).

[122]A proposition is a picture of reality. ... A proposition can be true or false only in virtue of being a picture of reality.

[123]Wittgenstein (1953, §136).

[124]And what a proposition is is in one sense determined by the rules of sentence formation (in English for example), and in another sense by the use of the sign in the language-game. And the use of the words "true" and "false" may be among the constituent parts of this game; ...

Another point is that instead of considering broader language games, i.e. linguistic activities other than asserting, we can, vice versa, consider narrower ones, i.e. to exclude even some kinds of assertions from consideration. It is, e.g., usually being argued that to consider the meaning of a statement as the class of possible worlds of which the statement is true is inadequate due to the statements reporting propositional attitudes such as beliefs: but it may be reasonable to argue that assertions about such attitudes belong to a "higher" language game than the "usual" one is. Similarly we can consider Frege's original proposal to consider meaning of a statement as its truth value as resulting from the assumption that what we nowadays call modal contexts do not belong to the basic assertive language game. In this way we can reach a kind of hierarchy of meanings (corresponding to a hierarchy of assertive language-games)[125].

[125]Cf. Sgall and Peregrin (forthcoming).

9 Criterial Reconstruction

9.1 What is an Explication?

In the preceding chapters we have embraced the idea of logic as a criterial reconstruction of necessary truth; and we have seen that the concepts of reference and meaning can be grasped as tools for such a reconstruction. In Chapter 2, moreover, we indicated that grammar can be approached analogously as a criterial reconstruction of "well-formedness"[126]. Now it is time to say more about the very concept of criterial reconstruction.

The general way of explaining something is to show that this something is something else, to show that something obscure is in fact in some way equivalent to something more transparent. In the most simple case the explanation is of the form *A is B*, where *B* stands for the transparent *explicatum* and *A* for the obscure *explicandum*. We can, for example, say that the largest city of the world is Mexico City - i.e., to reveal that the place pointed at by the concept of the largest city of the world coincides with the capital of Mexico. Or we may say that electricity is the motion of electrons - this allows the listener to see that the obscure force that manifests itself by moving elevators and lighting bulbs in fact coincides with an easily visualizable, mechanical process.

Both these cases of explanation, despite their different character, amount to coincidence 'in fact' - to coincidence that is in some sense casual and that is revealed by inquiring into the way the world actually is.[127] Thus, they are not a theme accessible to philosophers - they are the business of natural scientists. The philosopher is restricted to cases of *A is B* which are independent of how

---

[126]I.e. as the criterial reconstruction of the class of well-formed expressions, or of well-formed statements.

[127]It may be argued that the coincidence of electricity with motion of electrons is not casual. However, as long as we do not perceive the concept of electricity as *definitionally* equal to a way of electrons' motion, as long as we see it as an obscure power known only by its magnificent effects, there is something, namely just this obscure power, that may get revealed as casually coinciding with the electrons' motion.

the world actually is, and which are in this sense 'tautological'. A trivial example of such a 'tautological explanation' would be saying that electricity is electricity, or that necessary truth is necessary truth (cf. Section 4.8) However, the philosopher is surely not after platitudes of this kind - and he therefore faces the problem of articulating claims that are both 'tautological' and nontrivial.[128]

In his book, Carnap (1950a) proposed that the explanation a philosopher should offer would consist of the transformation of an imprecise, pre-scientific notion into a precise, scientific concept; he proposed to use the term *explication*. Developing this idea, Quine (1960a, pp.258-9) characterizes the enterprise of explication as follows:

> We fix on the particular functions of the unclear expression that make it worth troubling about, and then devise a substitute, clear and couched in terms of our liking, that fills those functions.

Quine's favourite example, the classical set-theoretical rendering of the concept of ordered pair, indicates that the paradigmatic case of explication is reconstruction within the framework of set theory.

This suggests a general answer to the question about the nature of philosophical explanation: doing philosophy - at least that kind of philosophy which we are pursuing here - means constructing comprehensible theoretical entities which are to reconstruct their pre-theoretical, incomprehensible precursors. And one possible way to understand comprehensibility in this context rests on the concept of *criterion* - comprehensible is what is criterial, because the criterion allows for a finite grasp on the infinite.[129]

Hence our proposal is that philosophical explication means stating the coincidence of two concepts, one pre-theoretical and therefore uncriterial, and the other theoretical and criterial. A typical example is the axiomatic method.

---

[128]The classical, Kantian solution to this problem was to make the difference between the *analytic* and the *a priori*: the claims of philosophy are 'tautological' in the sense that they are *a priori*, they are, nevertheless, in the same time nontrivial in that they are synthetic. A different answer was offered by Wittgenstein: what philosophy offers are, according to him, not real theses, but rather only hints that are like a ladder which is to be climbed and then kicked away. See also Section 10.7.

[129]Note the role of the concept of *criterium* in Ancient philosophy - see, e.g. Ros (1989, Buch II, §1.3).

It defines the concept of being provable that purports to coincide with the pre-theoretical ("intuitive") notion of being necessarily true. The claim *necessary truth is provability* is 'tautological' in that it conveys no information about the world (at least not in the sense in which *Electricity is a stream of electrons* does); however, it allows us to replace the uncriterial concept of necessary truth by the criterial concept of *provability* - we implicitly know what is necessarily true, but we can mechanically compute what is provable.

The problem of criterial reconstruction is closely connected with the concept of infinity; so let us start from the analysis of infinity.

9.2 Potential vs. Actual Infinity

We are so accustomed to viewing our world through the glasses of modern mathematics based on the Cantorian set theory that we usually take for granted that our world is full of infinite sets. We have an infinity of numbers, our language consists of an infinite number of expressions, and any physical body consists of an infinite number of space-time points. However, on second thought it is far from clear whether this view really is the most natural. On the contrary: bracketing our preconceptions, it seems to be quite plausible to conclude that we can ("really") be confronted (at once, but also during the whole finite time of our life) with at most a finite number of items[130] and that there hence is no real infinity present within our world.

What indisputably *is* present within the world is unlimitedness: the possibility to repeat, to continue various processes with no fixed limit. There is no fixed limit to counting, to creating new expressions, nor to imagining things divided into smaller parts; we can always conceive of a natural number greater than, an expression longer than, and a part smaller than, any given bound. But if the steps of a process cannot be limited by a bound, then it seems natural to call the process in-finite; moreover, it also seems natural to consider the steps of such a process, or whatever entities might be brought to life by the steps, as in-finite in number. So even on this view it is quite plausible to

[130]That we cannot be confronted with "all" numbers or "all" expressions of language is clear. We may be confronted with a thing said to consist of an infinite number of points; however, this clearly is not the confrontation with the infinity of the points.

ascertain infinity present within the world; however, not the actual infinity of the Cantorian set theory, but rather the potential infinity amounting to our expectance of the unlimited continuability of flows of actions or events.

From this point of view, the infinity of numbers consists in the unlimited possibility of counting, the infinity of language in the unlimited possibility of building new expressions, and the infinity of points of a thing in the unlimited possibility of thinking things divided into smaller parts. Acknowledging an infinite set thus means acknowledging some method (or methods) with the help of which we can always obtain new elements (possibly with the help of those obtained so far).

That the potential character of infinity comes naturally to human mind uncontaminated by modern mathematics can be documented by writings of almost all significant philosophers - since Antiquity till the outburst of modern mathematics - who did address the phenomenon of infinity. The ancient approach to infinity manifests itself clearly in the writings of Aristotle. Thus in *Book III* of *Physics* we can read[131]

> It is clear that the infinite takes different forms, as in time, in generations of men, and in the division of magnitudes. For, generally speaking, the infinite is so in the sense that it is again and again taking on something more, this something being always finite, but different every time. In the case of magnitudes, however, what is taken on during the process stays there; whereas in the case of time and generations of men it passes away, but so that the source of supply never gives out.

A similar attitude can be found within the writings of the pioneers of modern philosophical thinking. Thus, Locke (1690), claims that

> ... space, duration, and number, being capable of increase by repetition, leave in the mind an idea of an endless room for more; nor can we conceive any where a stop to a farther addition of progression; and so those ideas alone lead our minds towards the thought of infinity.

[131]Ross (1930, 206^{a}25, 30).

In our century, the situation changes. The almost universal acceptance of the Cantorian mathematics made the majority of scholars believe that actual infinity is something indispensable and straightforward; it made them, and consequently almost all of us, *see* the world in terms of actual infinity. However, there are still mathematicians and philosophers who keep reminding us about the actual infinity being secondary to the potential one - from Hilbert and Brouwer to Lorenzen (1957), Robinson (1965) and Vopěnka (1979). Perhaps the most illuminating analysis of the real character of the concept of infinity can be found scattered in the notes of Wittgenstein. Speaking about the Cantorian set theory as about "the theory of aggregates" he describes the situation as follows[132]:

> Die Theorie der Aggregate sucht das Unendliche auf eine allgemeinere Art zu fassen als die Theorie der Vorschriften. Sie sagt, daß das wirklich Unendliche mit dem arithmetischen Symbolismus überhaupt nicht zu fassen ist und daß es also nur beschrieben und nicht dargestellt werden kann. Die Beschreibung würde es etwa so erfassen wie man eine Menge Dinge, die man nicht in den Händen halten kann, in einer Kiste vepackt trägt. Sie sind dann unsichtbar, und doch wissen wir, daß wir sie tragen (sozusagen indirekt). Die Theorie der Aggregate kauft gleichsam die Katze im Sack. Soll sich's das Unendliche in dieser Kiste einrichten, wie es will.[133]

The point is to challenge considering infinity as something actual, but overreaching our finite possibilities restricting us to the potential infinity. If the only infinity we are really able to encounter in our world is the potential one, where then is the other one? As Hilbert (1925, p.190) puts it

---

[132]Wittgenstein (1964, §170). Wittgenstein's analysis of the infinite has been pointed out to the author by Wolfgang Kienzler.

[133]The theory of aggregates attempts to grasp the infinite at a more general level than a theory of rules. It says that you can't grasp the actual infinite by means of arithmetical symbolism at all and that therefore it can only be described and not represented. The description would encompass it in something like the way in which you carry a number of things that you can't hold in your hands by packing them in a box. They are invisible but we still know we are carrying them (so to speak, indirectly). The theory of aggregates buys a pig in a poke. Let the infinite accommodate itself in this box as best it can.

...das Unendliche findet sich nirgends realisiert, es ist weder in der Natur vorhanden, noch als Grundlage in unserem verstandesmäßigen Denken zulässig.[134]

Or, in an aphorism of Wittgenstein (1964, §123)

Es gibt keinen Weg zur Unendlichkeit, *auch nicht den endlosen.*[135]

9.3 The Infinite in Logic and Mathematics

Claims are often made that although the actual infinite is absent from the most natural picture of our world, it must be invoked if we want to develop an adequate logic and mathematics. But even this justification of the Cantorian prism is open to discussion.

It is clear that if we eschew actual infinity, then we would also be forced to eschew substantial parts of modern logic and mathematics. It is, however, not so clear that these parts would be really indispensable. Let us discuss the situation considering the example of the basic results of model theory of the first-order propositional calculus; for a detailed exposition of the range of problems, see Hailperin (1992).

It may seem that if we avoid actual infinity, then we are unable to talk about a model theory at all. Do not infinite domains lie in the very foundations of the theory? And are not the most basic results of the theory, such as Gödel's completeness theorem or the theorem of Löwenheim and Skolem, results concerning infinite sets?

In fact, theorems amounting to infinity must be divided into two groups. First, there are theorems, like that of Löwenheim and Skolem, which are directly about actual infinity and which would cease to make sense as soon as

[134]...nowhere is the infinite realized; it is neither present in the nature nor admissible as a foundation in our rational thinking.

[135]There is no path to infinity, *not even an endless one.*

we decided to eschew it. Such results surely cannot be used to substantiate the idea of actual infinity. Nobody convinced that there is "really" no actual infinity (let alone various different actual infinities), would miss them.

Then there are claims which seem to need actual infinity to be articulated or to be proven, which we, nevertheless, feel as significant independently of one's views of the actual infinity. This is, e.g., the case of Gödel's completeness theorem. The theorem says that the formulas provable within the first-order predicate calculus are just those which are true in every "reasonable" model structure; it thus says something that is not directly about actual infinite sets. The actual infinity is not the direct subject of the theorem, but rather only a tool used to establish the result.

However, to address completeness with the help of actual infinity is only *one of the possible ways* to address it. Another way is the way outlined by Herbrand and further elaborated by other theoreticians[136]. Herbrand avoids actual infinity completely and the results reached by his methods are not "worse" than those of Gödel; on the contrary, they may be considered to be essentially stronger in that they not only guarantee the existence of a model of every consistent first-order theory, but also show how such a model can be constructed[137].

The situation is similar with respect to other parts of logic based on the actual infinity, and with respect to mathematics. The vast theory of the relationships between various levels of infinity is respectable; it is, however, not clear if it really has any significance outside of mathematics. As Vopěnka (1979, p.9) puts it, what is intended to be mathematics *based on* the Cantor set theory is more and more only mathematics *of* the Cantor set theory.

---

[136]Cf. Shoenfield (1967).

[137]See Hailperin (1992).

9.4 Constructional Systems

All of this indicates that it is natural to view actual infinity as a mere correlate of the potential infinity[138]. Such a view then yields many interesting consequences, the majority of which are beyond the scope of the present essay. What we want to show is how this view can be understood to underlie the notion of criterial reconstruction leading to positing entities such as meanings. But to be able to do this we must first articulate the whole situation with some rigor.

We can consider a *constructional system* to consist of a finite number of items, called *basic (or trivial) constructs* of the system, and a finite number of operations, called *basic constructions* of the system. Each of the constructions is of a fixed arity and when applied to the appropriate number of constructs it yields either a new construct, or nothing at all. We shall say that something is a *construct* of a given constructional system either if it is a basic construct of the system, or if it is yielded by a basic construction of the system, i.e. if it is the value of the construction for some constructs of the system.

The simplest constructional systems are those the constructs of which are letters and concatenations of letters; these systems can be directly exhibited in writing. To present some examples, let us introduce a terminology for the description of ways of concatenation: the concatenation of two letter strings will be called the *suffixation* of the second to the first, or the *prefixation* of the first to the second. The concatenation of three strings will be called the *infixation* of the middle one to the outer ones. Thus the string **ABC** may be said either to be the suffixation of **C** to **AB**, or the prefixation of **A** to **BC**, or the infixation of **B** to **A** and **C**.

If we take the digit **0** as the only basic construct and the suffixation with **'** as the only basic construction, we have the constructional system N of basic numerals. The constructs of the set are the strings consisting of **0** followed by

[138]The traditional defendants of the primacy of potential infinity would insist that potential infinity is what is really *given* to us, what we really perceive; and that actual infinity is what is *inferred*, what we construe out of what we perceive. However, I doubt that a dividing line between the given and the inferred can be really drawn; so I think that the problem is not one of the confrontation of two concepts, but rather that of the confrontation of two ways of viewing the world, one engaging actual infinity and the second not.

any number of '. If we take instead the digits **0,1,2,3,4,5,6,7,8,9** as basic constructs and suffixation by any of the same digits as basic constructions, we gain the constructional system *DN* of decimal numerals.

Taking some letters or strings of letters as basic constructs and considering the basic constructions of the prefixation by $\neg$ and of the infixation by $\wedge$, $\vee$, and $\rightarrow$[139] we have the constructional system *PC* of classical propositional formulas. If we add the basic construction of prefixation by $\Box$, we gain the constructional system of modal propositional formulas.

More generally, we can accept that having a constructional system we can use its constructs to define a further ("higher-level") constructional system. The basic constructs of a system thus need not be listed, but can instead be specified via some other constructional system or systems; also the basic constructions of the system may be described with the help of other constructional systems. The only restriction is that circularity must be avoided.

So if we take the constructs of *DN* as basic constructs of a new constructional system and infixation with + and with • as its basic constructions, we have the constructional system *NE* (of numerical expressions). By taking the suffixation of a construct of *DN* to a letter as the basic construction we can gain constructional systems with the constructs $\mathbf{P}_1$, $\mathbf{P}_2$,...; $\mathbf{T}_1$, $\mathbf{T}_2$,...; etc. These constructs can be then used as *predicate letters*, *term letters* etc. to underlie the constructional systems of the formulas of various logical calculi, such as those of the language of the first-order predicate calculus (the atomic formulas of the language being constructed as the suffixation of a bracketed list of n term letters to an n-ary predicate letter; the formulas can then be understood as having been constructed on the basis of the atomic formulas by means of prefixation by $\neg$, infixation by $\wedge$, $\vee$, $\rightarrow$, and prefixation of the concatenation of $\forall$ with a variable letter).

Besides such constructional systems of letters we can also consider constructional systems of other kinds of objects. But these are essentially different in that they cannot be presented directly in writing: the only way to present them is to consider letters and strings of letters to act as symbols. (We may "define" a system of stone heaps by showing that some stones are "basic heaps" and by acknowledging that by putting two heaps together we "construct" again a heap. But heaps of stones cannot themselves be part of a theory; they

[139]Here and in some of the later examples we would have to redefine infixation so that it involves bracketing of the whole - but this is surely no substantial problem.

must be "stood for" by symbols. To write about a system of heaps, to make, so to say, a theory of the system, is thus possible only insofar as it is possible to consider some strings of letters to "stand for" heaps.)

If we consider some strings of letters to represent objects, then a set of such symbols can be considered to stand for the set of the corresponding objects. If we have a constructional system of strings of letters, we can consider it a system of *expressions*. To be able to identify the particular object any expression stands for, it is enough to be able to identify the object represented by every basic expression and to be able to identify the object represented by the result of every basic construction provided one can identify the objects represented by the expressions entering the construction. In fact, we need not be able to identify which particular objects individual expressions represent, it may be enough to say which pairs of expressions represent the same objects, and which pairs represent different objects. This is to say that we can make strings of letters into expressions by introducing a relation of equivalence between them. Thus we may constitute a constructional system of objects as a constructional system of letters plus an equivalence between the constructs of the symbol system.

We can consider the members of *N* to stand for *numbers* in the one-to-one manner; *N* can be considered hence not only as the system of *numerals* (its constructs being strings of digits), but also as the system of *numbers* (its constructs being denotations of the strings of digits). The constructs of *NE* can also be considered as standing for numbers; but in this case not in a one-to-one manner. Which two constructs of *NE* denote the same number and which do not is the matter of arithmetic. Similarly we may say that the formulas of the propositional or predicate calculus stand for propositions (whatever these may be considered to be); to determine which two formulas denote the same proposition, i.e. which two formulas are equivalent, is the matter of the logical calculus in question.

Also every finite set can be considered as a trivial constructional system all the constructs of which are basic. So we can understand the set $B = \{\mathbf{t},\mathbf{f}\}$ of the two truth value symbols as a constructional system. And in view of what has been just stated, we can consider B as not only the set of the two truth value symbols, but also directly as the set of the two truth values.

9.5 General Constructions and Functions

We may consider mappings of constructs of a system onto those of another system; but as we understood the range of the constructs of a system as given only via the system itself, we have to consider the mappings of constructs as given only via some kind of mappings of the systems. To be able to articulate the notion of mapping of constructional systems we must first introduce the general concept of construction.

The basic constructions of a constructional system are not the only way to obtain "new" constructs from "old" ones. We can gain new constructions by composing the existing ones. Besides this we can consider some general trivial constructions, such as the choice of one of an n-tuple of elements or "constructing" an element which was already constructed before. Let us define rigorously the general notion of construction. The concept of *(general) construction* of a constructional system S is defined by (1)-(5)[140]:

(1) *[basic construction]* Every basic construction of S is a construction of S.

(2) *[constant "construction"]* If c is a fixed construct of S, then the operation which yields c constantly for every $c_1,\ldots,c_n$ is a construction of S.

---

[140]Let us note that the concept of *construction* is central to the recent book of Tichý (1988) - the author claims that it is in general constructions that get expressed by language expressions. Note, however, that Tichý's constructions are something quite different from constructions as understood here: whereas we take constructions to be simply certain kinds of functions, Tichý's constructions should be rather understood as something like individual instances of application of functions to arguments. This means that whereas we take, e.g., the function + directly as a construction, for Tichý there is the construction of applying + to 1 and 2, there is another construction of applying + to 2 and 3, etc. Tichý's approach thus transcends traditional mathematics, which does not provide for understanding of a particular instance of application of a function to arguments as an object. In connection with the application of + to 1 and 2 we are normally able to speak of four objects: the function (+), the arguments (1 and 2) and the value resulting from the application (3); but there is assumed to be no such object as the instance of the application. We could, of course, model the instance as, e.g., the ordered triple $<+,1,2>$ - but Tichý's point is that these instances are something that should be seen as lying at the very heart of mathematics and of the theory of language.

(3) *[projective "construction"]* The operation which yields c_i for every $c_1,\ldots,c_n$ (for some fixed $i\in\{1,\ldots,n\}$) is a construction of S.

(4) *[composition of constructions]* if C is an m-ary construction of S and $C_1,\ldots,C_m$ are n-ary constructions of S, then the operation which for every $c_1,\ldots,c_n$ yields $C(C_1(c_1,\ldots,c_n),\ldots,C_m(c_1,\ldots,c_n))$ is a construction of S.[141]

(5) *[finite modification of a construction]* An operation which differs from a construction of S only for a finite number of arguments is a construction of S.

(6) Nothing else than what is specified by (1)-(5) is a construction of S.[142]

Let us consider some examples of nonbasic constructions. *NE* contains the basic construction which for N_1 and N_2 yields $N_1 + N_2$. The 3-ary operation which yields $N_1 + N_2 + N_3$ for N_1, N_2 and N_3 is a construction (it is the composition of the basic construction of the infixation of + with itself); so too is the unary operation which yields $N_2 + N_2$ for N_2 (it is the composition of the infixation of + with two unary projections). As every operation over a finite set is, according to (5), a construction, every truth-function is a construction of *B*.

---

[141](1) and (4) can be replaced by the single clause: if C is an m-ary basic construction of S and $C_1,\ldots,C_m$ are n-ary constructions of S, then the operation which for every $c_1,\ldots,c_n$ yields $C(C_1(c_1,\ldots,c_n),\ldots,C_m(c_1,\ldots,c_n))$ is a construction of S.

[142]If we considered a constructional system as an algebra (all its constructs constituting the carrier of the algebra and its basic constructions constituting the operations of the algebra), then the notion of general construction would be close to the notion of a polynomial over the algebra (the rule 5 is what makes a difference). For the notion of polynomial see Janssen (1983). Let us note that constructions defined in this way are not the only operations which may be considered in connection with a constructional system. We may moreover consider operations which can be called *quasiconstructions*. Quasiconstructions arise out of the decomposition of constructions: if an n-ary construction C yields c for some $c_1,\ldots,c_n$ and if $c_{i1},\ldots,c_{ik}$ are some of $c_1,\ldots,c_n$, then we may consider a k-ary quasiconstruction which applied to $c_{i1},\ldots,c_{ik}$ yields an intermediary construct which can then by a construction be further combined with the remaining of $c_1,\ldots,c_n$ to yield c. If general constructions are what corresponds to introducing polynomial symbols, quasiconstructions can be perceived as that which corresponds to introducing lambda-abstraction.

Let now S_1 and S_2 be constructional systems. A *mapping from S_1 into S_2* consists of (i) an assignment of a construct of S_2 to every basic construct of S_1; plus (ii) an assignment of an n-ary construction of S_2 to every basic n-ary construction of S_1. If m is a mapping from S_1 into S_2 then the value m(c) for a construct c of S_1 is defined as follows: (i) if c is basic, then m(c) is given directly by the definition; (ii) if $c=C(c_1,\ldots,c_n)$, then $m(c)=C'(m(c_1),\ldots,m(c_n))$, where C' is the construction assigned to C.

A typical example of a mapping from a constructional system is the truth valuation of propositional formulas understood as a mapping from *PC* into *B*: basic constructs (atomic propositions) of *PC* are mapped on arbitrary basic constructs (truth values) of *B*; the basic constructions of *PC* are mapped on the appropriate truth functions of *B*. Another simple example is a mapping from *PC* into *N* expressing the rank of a formula: every basic construct of *PC* is mapped on the basic construct **0** of *N* and every basic construction of *PC* is mapped on the only basic construction of *N*.

A mapping from a constructional system S into *B* is called *Boolean*[143]. A Boolean mapping from S means the classification of the constructs of S into two groups: into those which are mapped on **t** and those which are mapped on **f**. If we consider a constructional system as specifying a class of items (the class of all its constructs), then such a Boolean mapping can be considered as demarcating a subclass of this class.

The first of the examples of mappings of constructional systems we have given is a Boolean mapping; it can be understood as demarcating a subclass of the class of all the formulas of the propositional calculus. The class that is so demarcated is nothing else than a consistent theory.

9.6 Constructional Systems and Sets

Having a constructional system we can consider all its constructs to constitute a set. More precisely, we can consider two constructional systems to be equivalent if every construct of one of them is also a construct of the other and vice versa, and we can consider the set as that which all such equivalent

[143]Note that we use the term 'Boolean mapping' for a mapping whose range is the Boolean algebra of two values, not a more general Boolean algebra.

constructional systems share. (The relation *x is a construct of the constructional system S* is clearly congruent with respect to this equivalence, and hence it can be directly turned into the relation *x is an element of the set S*.) Hence sets can be considered as equivalence classes of constructional systems: the difference between a set and a constructional system is that the identity criterion for sets is the sameness of elements, while two constructional systems can be different despite having precisely the same constructs.

We can further consider parts of the ranges of a constructional system; we can, e.g., consider the set of all the even numbers. How can a part of a range of a system be selected? We have to say which items of the range do belong to it and which do not; which is to say that we have to define a Boolean mapping of the range. In our particular case, all the odd natural numbers can be considered as separated from *N* by means of the Boolean mapping b such that $b(\mathbf{0})=\mathbf{f}$ and $b(x')=\mathbf{t}$ iff $b(x)=\mathbf{f}$. A consistent theory can be separated from the range of all the formulas of the propositional calculus by means of the Boolean mapping described above. This can be called *direct separation.*

However, we must also acknowledge more intricate ways to separate a part of a range of a constructional system. Let us consider the constructional system of modal propositional formulas with the basic formulas F_1, F_2,... . As we shall se in the next paragraph, the consistent theory containing F_1, F_2, $\Box F_1$, but not $\Box F_2$, cannot be separated by means of a Boolean mapping; hence, if we insisted that the only way to extract a part of the range of a constructional system is direct separation, then we would lose the possibility of taking such a theory into account at all.

The reason why the theory in question is not directly separable is that if m were a Boolean mapping which would accomplish the separation, i.e. if m assigned $\mathbf{t}$ to P_1, P_2, $\Box P_1$ and $\mathbf{f}$ to $\Box P_2$, then $\mathbf{t}= m(\Box P_1)= m(O(P_1))= O'(m(P_1))= O'(\mathbf{t})= O'(m(P_2))= m(O(P_2))= m(\Box P_2)= \mathbf{f}$ (where O is the construction of prefixation of $\Box$ and O' is the corresponding construction of *B*). However, the theory can be separated in a more intricate way with the help of auxiliary items: using a set $\{a,b,c\}$ we can define a mapping m_1 from the formulas into $\{a,b,c\}$ and a mapping m_2 from $\{a,b,c\}$ into *B* such that the composition of m_1 and m_2 assigns $\mathbf{t}$ to P_1, P_2, $\Box P_1$ and $\mathbf{f}$ to $\Box P_2$. It is enough to let $m_1(P_1)= m_1(\Box P_1)= a$, $m_1(P_2)= b$, $m_1(\Box P_2)= c$ and $m_2(a)= m_2(b)= \mathbf{t}$, $m_2(c)= \mathbf{f}$. The composition is, however, not a mapping itself; and hence the separation is *indirect*.

So we can conclude: we can consider a set as either a range of a constructional system, or a part (directly or indirectly) separated out from such a range. In this way the traditional extensional notion of set can be reached in a way that seems to be adequate to its real foundation[144].

9.7 Criterialization and Posits

We have claimed that it is natural to consider infinity to be primarily grounded in an unlimited productive activity. We have articulated the notion of a constructional system which can be considered to theoretically reconstruct the relevant structure of such an activity. However, and this is the crucial point, not every unlimited production of new items can be primarily understood as a matter of such an explicit system.

The infinity of formulas of a formal language is primarily a matter of the unlimited applicability of some formal rules (this is the *primary* way in which such a system is given); the infinity of expressions of our natural language is, on the other hand, a matter of the human faculty to produce new expressions without a fixed limit. We may call the infinite sets of the first kind, as we have done in concrete cases before, *criterially infinite* sets; while calling the other *uncriterially infinite*[145].

---

[144]It is interesting to consider the question whether the roundabout is of some significance; whether the result is the same as if we took sets as primitives. Any range of a constructional system is clearly a set in the traditional sense; however, is, vice versa, anything that counts as a set in the traditional sense a range of a constructional system? It is clear that by way of constructional systems we can reach an at most denumerable number of sets; whereas there are more sets in the Cantorian universe. However, the question which traditional sets are so reachable is not easily answered, as the answer depends on what we are willing to accept as a construction.

[145]Barbara Partee has pointed out to me that the Chomskian view shared by many linguists would be that the mentioned human faculty is necessarily a manifestation of an explicit grammar contained within the human mind, but not available to conscious inspection. This claim seems to me to be either trivial, or highly controversial: if, on the one hand, we take saying that our minds contain a grammar as only a "theoretical way" of saying that we are apparently capable of using language, then it is trivial; if, on the other hand, we take it as saying something more, then the grammar comes to be an 'idle wheel', for it is something that exists independently of its overt manifestations, but cannot be found about independently

The criterially infinite sets are *comprehensible*, as they can be comprehended by means of comprehending their (finite) criteria; uncriterially infinite sets are, on the other hand, *incomprehensible*[146]. This is why we take pains to reconstruct uncriterially infinite sets as criterially infinite ones: we want the incomprehensible to become comprehensible. If there is an uncriterial set, we try to find a criterial one which would coincide with it (i.e. we try to find a constructional system which would construct the as yet incomprehensible set); only then can we consider the set as really (theoretically) understood. This is the ultimate idea lying behind the notion of *criterial reconstruction*.

There are uncriterially infinite sets which can plausibly be reconstructed as ranges of constructional systems; but there are also sets which can plausibly be reconstructed rather as parts of such ranges, which are to be separated from the ranges. We have seen that such a separation cannot always take place directly, that to accomplish it we must sometimes make use of auxiliary entities. Let us elucidate this once more with the help of a simple example.

Let us consider the constructional system of all possible constructs built of a given set of building-blocks. Let us assume that some of the constructs distinguish themselves as being inhabitable houses. Inhabitable houses can hardly be separated out from all the constructs directly: this would mean that whether a construct is an inhabitable house depends only on its parts being inhabitable houses. But this is hardly the case: to decide whether something is inhabitable we need to know more than if its parts are inhabitable. It is, however, possible to compositionally evaluate all the possible constructs in such a way that inhabitable houses obtain an evaluation different from that of other constructs.

Let us make the point clear with the help of a much greater oversimplification; let us assume that we build "houses" out of children's building blocks. Let us assume that we have the part-whole system of the

---

of them. So the latter claim seems to me to incorporate a pseudoexplanation similar to that used when we try to explain semantics in terms of metaphysically grasped meanings. As Pears (1994, p.48) puts it, we try to see the situation in terms of two levels of reality, a 'lower' level (the overt linguistic practises) and a 'higher' level (the grammar within the depths of human consciousness, or the meanings in the heights of a platonic heaven), "start with something substantial on the lower level, project its shadow onto a higher level and then take the shadow to explain the substance."

[146] Elsewhere (Peregrin, 1992) I have proposed to understand Wittgenstein's (1953) *Unübersehbarkeit* in this sense.

constructs depicted in Figure 3. 1,2 and 5 are simple buildings (the basic blocks), 3,4,6 and 7 are complex ones. The constructs 6 and 7 count as houses. Let us, moreover, assume that all the complex constructs arise out of their parts by the same construction ("putting one construct on top of another").

The class of houses cannot be separated out from the class of all constructs directly: a house may consist of two non-houses (6 of 3 and 5), but not everything that consists of two non-houses is a house (3). We can, however, evaluate the parts as *oblongs* (1,2), *squares* (3,4), *triangles* (5) and *houses* (6,7) and say that two oblongs add up to a square and that a triangle and a square add up to a house. The separation of houses out of all the constructs thus necessitates entities of a new kind, entities which could be called *geometrical figures*: *oblong*, *triangle* and *square*[147].

This way of positing new entities can be considered as constituting a general paradigm, which underlies the special cases of positing objects discussed in the previous chapters. Referents of expressions and meanings of expressions, but also possible worlds or situations or model structures, can be seen as posits brought into being by various kinds of (successful or unsuccessful) attempts at a criterial reconstruction of truth.

9.8 Infinity, Rules and Meanings

We have seen that infinity is insinuated by certain kinds of recursively applicable rules or by repeatable activities reconstructible by way of such rules. We may consider rules to be implicit to our usage of language, or even to other kinds of our activities; however, it is only formalization that makes the rules explicit. It is what we have called criterial reconstruction that makes the recursive character of our language visible. Formal logic and formal mathematics are primarily not something with which to refer to objects and their sets; they are, first and foremost, the explication of, as Skolem (1923) puts it, "die rekurrierende Denkweise", the recursive mode of thought.

We have seen that quantification is a matter of our tendency to consider certain grammatical construction as amounting to infinity. Now we see, more

---

[147]For a more detailed elaboration of these ideas see Peregrin (forthcoming b).

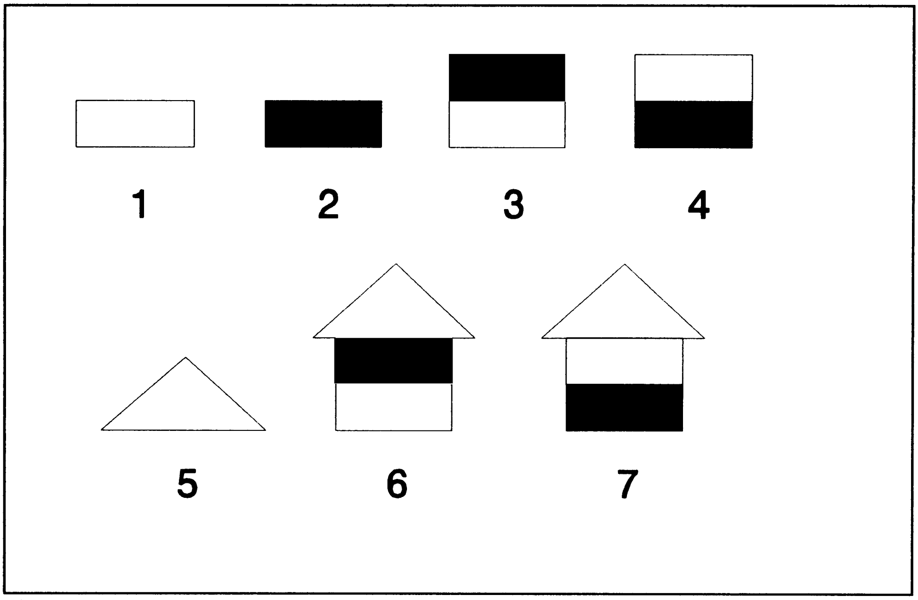

Figure 3

generally, that infinity is something that comes to mind when we manipulate expressions and signs in a certain way and which becomes explicit when we capture the manipulation by means of explicit, formal rules. In this sense, "we attain to the infinite by means of signs."[148] The concept of infinity thus can be seen as closely linked to the concept of rule.

Now the process in which we make the recursively applicable rules underlying our linguistic activities explicit, the process which can be considered as a process of criterial reconstruction of "uncriterially" infinite sets of expressions, is the process in which we also explicate the notion of meaning. Meaning is what we require for carrying out the reconstruction, what is necessitated by the reconstruction. Meanings, and more generally, values, are ways of viewing objects from the vantage point of their manipulation according to rules. Where there is a rule, there are ways of acting, of using items according to the rule; and there are items appearing to be usable in various ways. This means that where there are rules, there are equi-valences and

[148]Goldfarb (1977, p.361).

values; there are meanings. This indicates that the concept of meaning is closely allied to the concept of rule.

Finally, what makes meaning indispensable is the infinity of ways an expression can be used; the concept of meaning amounts to a finite, "packed" summarization of the behaviour of the expression in an infinite number of cases. This closes the circle and we can conclude that the concepts of infinity, rule and meaning are essentially interwoven.

10 Necessity

10.1 The Quinean Challenge

During the preceding chapters, we have been approaching logic as the pursuit of necessary truth. Moreover, we have seen also that semantics can be understood as a matter of the same pursuit - necessary truths amount to rules governing our usage of language (they can be seen as articulating instances of consequence and hence limitations to assertibility), and as expressions acquire meaning just via being manipulated according to these rules, it is necessary truth that underlies meaning. We have taken for granted that necessary truth exists, that statements of our language can be divided into those which only happen to have their truth value and those which have their truth value fixed once for all. From this viewpoint, the necessary/contingent distinction is thus absolutely crucial. It is the existence of this distinction that makes logic meaningful; and it is this distinction that makes it possible to make sense of the concept of meaning.

However, is the distinction really such a straightforward matter? Well, it seems to be quite clear that the statement such as *Fido is a dog* only *happens* to be true and is thus contingent; and that, on the other hand, the statement *If Fido is a dog, then Fido is a dog* evidently cannot be false and is thus necessary. However, what about such statements as *If Fido is a dog, then Fido has kidney*? Should we find out that Fido's inner workings differ from those of a normal animal in that he can live without kidney, would we conclude that he is not a dog? Or should his kidney be removed after his death, should we refrain from calling his corpse a dead dog?

We may assume that it is not necessary truth in general, but only logical truth that is of direct concern for logic. However, as can be seen, e.g., from the endless quarrels between intuitionists and the adherents of classical logic, even the delimitation of logical truth appears to be far from unequivocal; and there is little outlook that we are in any way converging towards a unique universally acceptable resolution.

This indicates that both the necessary/contingent and the logical/extralogical distinction are essentially vague. However, this is not the most disastrous challenge these distinctions have to face; the greatest trouble is constituted by the claim, articulated in the most popular way by Quine, that such distinctions do not make any real sense at all. In his famous paper, Quine (1951) argues that there are no statements which could be absolutely classified as necessary, that there rather are only statements whose large-scale infiltration to our discourse makes it troublesome to change their truth values. There is, according to Quine, nothing "in the nature of things" which would make statements contingent or necessary, there are only our pragmatic calculations based on the subconscious assessment of the upheaval the change of the truth value of a statement in question would cause within our overall picture of the world.

However, if the status of the necessary/contingent distinction were so uncertain, what about logic? If, as we have concluded, logic is a matter of the criterial reconstruction of necessary truth, would we not have to conclude that it is chasing a mere chimera? And what about semantics - does not the Quinean challenge question the very idea of meaning? These questions appear to be of the utmost importance; however, to be able to answer them properly, we must first investigate the nature of necessary truth in greater detail.

10.2 Words and Things

The necessary/contingent distinction has a history; the contemporary way of viewing it dates from the beginning of the modern era, when it took variously the shapes of Leibniz's differentiation between *truths of fact* and *truths of reason*, of Hume's between *matters of fact* and *relations of ideas*, and of Locke's between *real* and *verbal truth*, where real truth is "about ideas agreeing to things", whereas verbal truth is a matter of "terms being joined according to the agreement or disagreement of the ideas they stand for".

Kant, besides his proposal to the effect of splitting the distinction into the *a priori/a posteriori* and *analytic/synthetic* distinctions, showed that its significance is not only epistemological, but also ontological. He pointed out that the question about the character of the world cannot be separated from the question about the character of the way we apprehend the world, i.e. about the structure of our reason. This motif then reappeared with a vengeance in the present century, after the *linguistic turn*, as appearing in the writings of Russell, Carnap, Wittgenstein and others. These philosophers realized that the way our world is cannot be separated from the way we describe it; and hence that what Kant took to be the structure of our reason is in fact the structure of our language.[149]

The turn was often understood as inseparably connected with the *theory of correspondence*, as we discussed it in Chapter 6 - i.e. with the philosophical companion of the commonsense view that the purpose of individual expressions of language is to represent things (either directly, or via the capacities of human minds). According to this view, statements stand for some kind of junctions of things, and a statement is true if the things really are joined in the way the statement declares them to be. This view presupposes direct correspondence between words and things; and it thus suggests the view of the world as a basic supply of objects continually changing their configuration. Categories of objects, i.e. ontological categories, correspond to categories of expressions, i.e. to grammatical categories; true statements correspond to actually existing junctions of objects and false ones to junctions of objects that are possible, but not actual. This is the picture drawn by the logical atomists - by Russell, early Wittgenstein, and also by Carnap - as expounded in Section 5.6.

The theory of correspondence suggests the picture that contingent truths reflect the fluctuating "part" of the world and necessary ones the "part" that is fixed. Typical contingent truths, such as (32), reflect the (accidental) instantiation of a property by an individual (in the case of (32) of mortality by Aristotle); typical necessary truths, such as (33), on the other hand express the (eternal and inevitable) subordination of a property to another property - in the case of (33) of manhood to mortality.

Aristotle is mortal (32)
(Every) man is mortal (33)

---

[149]Kant's proposal to the effect of making the problem of the "synthetic *a priori*" into the crucial question of philosophy got abandoned during the turn - thus, although this problem is still imporatant to Frege, it is either not important, or completely alien to his followers, to Russell, Carnap and Wittgenstein.

It is then either possible to be overtly platonistic and to take the difference between instantiation and subordination as the mere matter of the fact that the former is only momentary, whereas the latter is eternal, i.e. to say that the subordination of manhood to mortality amounts to the constellation of properties in the same way as the instantiation amounts to the constellation of individuals and properties; or it is possible to consider the talk about subordination as only a metaphoric way of talking about limits of instantiation, to consider saying that manhood is subordinate to mortality as a way of saying that whatever instantiates manhood cannot fail to instantiate mortality. In the former case we take necessary truths as portraying junctions of objects in the same way as the contingent ones do, whereas in the latter case we take them to articulate limits to such junctions. Whichever of the two we choose, we have to see the boundary between necessary and contingent truth as grounded "in the nature of things": in the first case as the boundary between the steady and the fluctuating part of the world, in the second as the boundary between its *"form"* and its *"content"*.

10.3 The Dogma Rejected

It is this kind of representational and atomistic view of language which can be considered as underlying the views of Schlick, Carnap and other logical empiricists of the Vienna Circle. The central concept of their epistemology was *verification*: how to verify a statement is a matter of reason, whereas the verification itself is a matter of fact. The way in which a statement is verified is the key to its meaning: a statement whose way of verification is not settled is considered meaningless. In a programmatic article, Moritz Schlick (1932) proposed to distinguish between what he called *the pursuit of truth* and *the pursuit of meaning*: the scientific exploration of the content of the world, pursuing truth, must be preceded by the philosophical explication of the form of the world, pursuing meaning.

It is this picture that was challenged by Quine; and this challenge has become one of the most discussed, and also one of the most controversial philosophical doctrines of this century. Quine's point is that the boundary between necessary and contingent truth is not an absolute one, and instead only a more or less pragmatic matter. Thus, Quine (1963, p.37) claims

> a boundary between analytic and synthetic statements simply has not been drawn. That there is such a distinction to be drawn at all is an unempirical dogma of empiricists, a metaphysical article of faith.

Quine's point is that we can never verify a single isolated statement because only whole theories (or at least their nontrivial parts) can be really verified. A scientist makes an experiment to verify a hypothesis, but what he in fact verifies is the hypothesis plus a backlog of theory he takes for granted; and the fact that he considers one statement as hypothesis and all the others as the unquestioned backlog is a purely pragmatic matter. A statement appearing as necessarily true is thus, according to Quine, simply a statement which is so interwoven with the "web of our beliefs" that we cannot imagine giving it up without doing the whole web a serious harm. Hence (ibid., p.42)

> it is nonsense, and the root of much nonsense, to speak of a linguistic component and a factual component in the truth of any individual statement. Taken collectively, science has its double dependence upon language and experience; but this duality is not significantly traceable into the statements of science taken one by one.

Quine thus claims that no absolute boundary between analytic and contingent statements can be really drawn; not only because such a boundary would have to be fuzzy, but rather that it would be wholly arbitrary and hence makes no sense at all. What makes one statement more easily refutable than another is, according to Quine, not something in the nature of things, but rather only some kind of - usually subconscious - pragmatic calculation. If a hypothesis implied by a theory comes to be falsified, we must fix the theory to cancel the implication of the false hypothesis by the removal of some statement(s) of the theory; what Quine claims is that in such a case we are in principale free to remove *any* statement of the theory, that the only guiding principles are pragmatic principles of simplicity and austerity. He argues that there is nothing "in the nature of things" that would force us to keep one statement "come what may". The only difference is that the removal of some statements would cause only a minor further adjustment, while that of others, e.g. of the laws of mathematics, would bring about the necessity of complete readjustment of the rest of the theory.

This conclusion, and the convincing evidence Quine brings to its support, questions the very possibility of making sense of the distinction between necessary and contingent truth and of decomposing the content of an expression into two separate parts: into a fixed meaning and an alternating reference. It also questions the possibility of decomposing the world into a fixed form and an alternating content; which then leads directly to Quine's indeterminacy theses.

10.4 Wittgenstein on Rules

Quine's holism, however path-breaking a doctrine it is often considered to be, is in fact no real *novum*, for the same kind of insight has been reached and widely discussed earlier by Wittgenstein. Wittgenstein's teaching was essentially cryptic and it was Quine who reallly popularized the holistic insight; however, the depth of some of Wittgenstein's commentaries remains unmatched.

Wittgenstein's philosophical development since the thirties up to *Philosophical Investigations* stems from the recognition of the multiversity of activities in which our language engages. Thus, Wittgenstein (1953, §23) claims

> Wieviele Arten der Sätze gibt es aber? Etwa Behauptung, Frage, und Befehl? - Es gibt *unzählige* solcher Arten: unzählige verschiedene Arten der Verwendung alles dessen, was wir "Zeichen", "Worte", "Sätze", nennen.[150]

The subject of the Fregean logic as well as of Wittgenstein's own teaching of the *Tractatus* is asserting; however, as we have pointed out in Section 8.7, asserting is only one activity among many others. It is only one of

---

[150]But how many kinds of sentences are there? Say assertion, question, and command? There are *countless* kinds: countless different kinds of use of what we call "symbols", "words", "sentences".

the vast quantity of possible language games[151]. It is perhaps a game more important than others (it is this game that is primarily relevant with respect to the question *what is there*), but nevertheless it is not *the* game.

Now if there were only one game to be played with our language, then the rules of the game, its "logic", would be not "in the world", but rather transcendent to the world. However, there are, Wittgenstein insists, multifarious games; and to say this is to say that the rules *are* in the world. Our assertions may be understood as moves within an actual language game; however, we can also state rules of games (if rules are "in the world", they are no more unspeakable!). This is the source of an important ambiguity (ibid., §242):

> Zur Vesrtändigung durch die Sprache gehört nicht nur eine Übereinstimmung in den Definitionen, sondern (so seltsam dies klingen mag) eine Übereinstimmung in den Urteilen. ... Eines ist, die Meßmethode zu beschreiben, ein Anderes, Messungergebnisse zu finden un auszusprechen.[152]

The rules are one thing, and the moving according to these rules is another.

The rules of the "assertive game" are the necessarily true statements; these statements are what set up the boundaries of what can and what cannot be reasonably asserted. To accept statement (33) as a necessary truth means to prohibit the simultaneous assertion of *Aristotle is human* and *Aristotle is not mortal.* In *Tractatus* Wittgenstein had to take necessary statements for nonsensical: it does not make sense to say what is transcendent and what thus in fact *cannot* be said. In *Philosophical Investigations* the very concept of rule, and implicitly also that of necessary truth, is put into question.

The distinction between a rule of a game (a necessary statement, or, as Wittgenstein calls it, a "grammatical" statement) and a move in the game (a contingent assertion) is thus central to *Philosophical Investigations*. However, the difference and its relativity was first pointed out by Wittgenstein much

---

[151]And it was awareness of the *uncountability* of the games which prevented Wittgenstein from developing an Austinian theory of speech acts.

[152]If language is to be a means of communication there must be agreement not only in definitions but also (queer as this may sound) in judgements. ... It is one thing to describe methods of measurement, and another to obtain and state results of measurements.

earlier, in a notably illuminating way in his Cambridge lectures published as *The Yellow Book* (see Ambrose, 1979).

10.5 The Arbitrariness of the Possible

In those lectures Wittgenstein took pains to clarify the proper distinction between theories as such and hypotheses within such theories. Hypotheses, according to him, can be true or false, but theories themselves can at most be practical or impractical. As Wittgenstein (Ambrose, 1979, p.69) puts it

> Some theories are practical and some impractical. Impractical systems are rejected, and the rejection is treated as though what is rejected is something false.

To understand Wittgenstein properly we must understand what *theory* means for him in this context. A theory in Wittgenstein's sense is not a class of statements, as the term is sometimes understood within modern logic; a theory in his sense is constituted by the rules of the language game to be played. And to say that it is constituted by the rules is to say that it is in a sense wholly arbitrary: the rules are unsurpassable as long as we are inside of the game, but they appear arbitrary as soon as we move outside of it.

In fact, any statement can be considered either a rule or a hypothesis. This is what Wittgenstein (ibid., p.70) illustrates by a remarkable example[153]:

> Suppose that a planet which according to a certain hypothesis describes an ellipse does in fact not do so. We should then say that there must be another planet, unseen, acting on it. It is arbitrary whether we say our laws of orbit are right, that we merely do not see the planet acting on it, or that they are wrong. Here we have a transition between a hypothesis and a grammatical rule. If we say that whatever observation we make there is a planet nearby, we are laying this down as a rule of grammar; it describes no experience.

[153]See the discussion of this example by Leilich (1983).

This illustrates the reason underlying the Quinean rejection of the absoluteness of the analytic/contingent opposition more succintly than any of Quine's own examples. We have a predictive statement *the planet describes such-and-such ellipse* implied by a theory and we find out that the hypothesis is false. We have to amend the underlying theory by retracting some of its assumptions; we may either reject the tacit contingent assumption *there is no planet nearby*, or we can abandon some of the "laws" of the theory, i.e. some of the statements posed as analytic by the theory. Wittgenstein clearly sees, as later Quine too, that there is no absolute reason to favour one of the two solutions and that there may be at most some relative reasons connected with the simplicity and elegance of the theory in question.

This implies that, as Quine (1963, p.43) puts it, "any sentence can be held true come what may, if we make drastic enough adjustments elsewhere in the system". But this implies also that, as Wittgenstein (ibid., p.69) puts it, "it is in a sense arbitrary what is called possible and what is not called possible.". Indeed, "sentences which by present conventions clearly make no sense, ..., can of course be given a sense. It is because whatever is said can be given sense that the conventions adopted are called arbitrary." (ibid., p.66). Rules are in this clear sense arbitrary; and so are analytic truths as articulations of rules. Hence the arbitrariness of what is necessary, and consequently of what is possible.

10.6 The Nature of Necessary Truth

To accept a language means to approve some of its statements as necessarily true. There is no need in principle to adopt a particular language (although our native language is in a sense forced upon us and it is just this language which furnished us with our ultimate means of theoretical coping with the world); however, doing theory (in the broadest sense of the word) and engaging oneself in rational argumentation presupposes adoption of a language as an unquestionable basis. If I ask *Is the statement (33) necessarily true?*, the straightforward answer is, of course, *Yes*. However, it is equally straightforwardly true that such an answer rests on the tacit assumption that (33) is the familiar statement of our shared language - we can surely imagine a language in which (33) would be false. There are good reasons for the assumption to "normally" remain tacit (in fact, this is what makes "normal" communication possible); but if we turn to philosophical questions, such as the question about the nature of necessary truth, we can no longer ignore it. And taking it at face value we realize that language is a Janus-faced being: its rules, its necessary truths, when seen "from outside", are contingent and arbitrary, but when seen "from inside", are necessary and obligatory.

To say that a contingent statement is true because of the way the world is may be misleading (it may lead us to the conviction that in paraphrasing the statement (32) as *The individual Aristotle instantiates the property of mortality* we explain the truth of the former), but it surely does make sense; it amounts to the evident fact that the truth value of such a statement is "supplied by the world" and that it can be discovered by the empirical investigation of the world. However, to say that a necessary statement is true because of the way the world is does not even make this kind of sense - establishing necessary truth is not a matter of investigation of the world, but rather of setting up the language.

To say that something is necessarily true is to declare one's conformity with a certain language game, one's willingness to accept a certain language. If I declare (33) to be necessarily true, then what I say is that were I to encounter an immortal being, I would not consider it a man[154]. If someone disagrees, then our disagreement cannot be settled by investigation of the world; his disagreement simply means that he is playing a language game different from mine. I can persuade him (for example by showing him some books from which it would follow that my game is the one played by the majority of speakers of English), but I cannot show that he is false in the same sense in which I could were he to claim that Aristotle is not mortal.

From this point of view, Quine's rejection of the meaningfulness of the concept of analyticity seems to be needlessly dramatic[155]. If I want other people to understand me and if I want to be able to argue with them, I have to abide a language game, or a conceptual scheme (or whatever one chooses to call it); I must succumb to certain necessary truths however arbitrary they may

[154]Hence Wittgenstein's (1956, §III.30) "das *Muß* einem Gleise entspricht, das ich in der Sprache lege" ["the *must* corresponds to a track which I lay down in language"].

[155]In fact, Quine himself seems to be no longer happy with his original strong formulation of his challenge to the analytic/contingent boundary - witness Quine (1991).

have appeared beforehand[156]. It is true, as Davidson (1974) warns us, that it may be deeply misguiding to view real languages or "conceptual schemes" as something of the kind of set-theoretically interpreted calculi of formal logic; but this does not mean that there would be nothing definite which I must obey if I want to be understood.

It is, indeed, in a sense arbitrary which statements are necessarily true - but only if we view language from outside. Accepting a language as our means of communication - and we need a language all the time - we relinquish our possibility of questioning its necessary truths, on pain of blurring the boundary between consistency and inconsistency and so losing the firm ground beneath our feet, downgrading language to a bundle of expressive shrieks.[157]

We can say that *necessary truths are what certain contingent truths, namely truths about the usage of language, become when they are seen from inside*. Considering the case of the necessary truth *2+2=4* Malcolm (1940, pp.198-199) writes

> What *justifies* you in putting down 2+2=4 in your calculation is a certain fact about the way everybody uses the expressions "2+2" and "4". But you do not *assert* this fact when in doing the calculation you write down or mutter to yourself "2+2=4". In other words, in these circumstances you are not using "2+2=4" to *describe a fact*; you are letting it play a role in a calculation.

---

[156]Stekeler-Weithofer (1986b, p.10) writes: "Viele Auskünfte in Wörterbüchern und Grammatiken sind zunächst *'deskriptiv'* zu verstehen, sie *beschreiben* den üblichen Sprachgebrauch. Aber sie werden zu *normativen Regeln* für den, der in der betreffenden Sprache etwas ausdrücken will: Wer verstanden werden will, hat in der Regel so zu sprechen, wie 'alle' zu reden pflegen, und er hat sich dabei in der Regel auf ein gemeinsames 'Grundwissen' zu stützen." ["Information given in dictionaries and grammars are, at first, descriptive. They say how one ordinarily uses the language. But they turn in to normative rules for anybody who wants to express something in that language. If I want to be understood I have to speak in a way everybody is used to. And I rely on common knowledge or common belief."]

[157]Thus, as Penco (1994, p.132) claims: "What is necessary is not what is arbitrarily held necessary, but what is accepted as necessary in response to our basic constitution and our needs."

We are inclined to think that a statement like (33) or $2+2=4$ is *either* a contingent truth amounting to the way we use language, *or* a necessary truth - however, in fact it is in a sense *both*, depending on the perspective we adopt. Being inside language we let such statements play their role in communication and we thus make them turn their necessary face to us; being outside we perceive their contingent face.

The obvious objection to such a relativizing view is the following: we can do logic and philosophy only within a framework of a fixed language. This claim is quite plausible if we understand *language* in the commonsense, relaxed sense: it is quite reasonable to avoid oscillating between, say, Czech and English. But if we understand language as that which fixes necessary truth, then the requirement of keeping the language fixed would not allow us to analyze just the philosophically interesting problems, and it downgrades philosophy to a set of platitudes taken for ultimate explanations. Philosophically interesting problem arise when we realize that there is some leeway for tampering with necessary truth without breaking down rational communication; and that *language is in fact nothing else than an equilibrium between the stability guaranteeing ongoing understanding and the variability making language into something more than a mere set of codes.*

10.7 Philosophy as the Pursuit of Meaning

What does this imply with respect to the nature of philosophy? Let us return to Schlick's proposal to distinguish between *the pursuit of truth* (yielding contingent theses "about the world") and *the pursuit of meaning* (yielding necessary interconnections between concepts we use). The aim of philosophy was thought to be simply to pursue and to bag meaning, leaving the natural sciences a free hand to pursue truth. In this way, the philosophers of the linguistic turn thought they had found a new definition of philosophy making it into a scientifically respectable enterprise and blocking attempts to make it produce what they considered unscientific metaphysical rubbish.

But Schlick's proposal, as is readily seen, is tricky. Does *the pursuit of meaning* mean the same as *the pursuit of truth about meaning*? It seems that in whichever way we answer this question, we are in trouble. If we do identify *the pursuit of meaning* with *the pursuit of truth about meaning*, then philosophy turns to be identical with a branch of science, namely empirical semantics; and if we reject the identification, then we are left with the conclusion that philosophy is not to yield truths, which seems to be absurd. In the former case we would have no philosophy at all, in the latter philosophy would be something mysterious and there could surely be no philosophical *theory*.

Logical positivists do not seem to have realized the full significance of this dilemma; they wanted to have philosophy as a theoretical discipline *par excellence* and so embraced the first answer. Thus Carnap (1934) considers meaningful philosophical statements simply as "second-order" contingent truths, as truths about the language in which we formulate the "first-order" truths about the world; philosophy, according to him, is "the logical syntax of the language of science". This seems to mean that philosophy is simply a science, a particular brand of linguistics; but Carnap insists that philosophy, at the same time, constitutes the *foundation* to science.

Carnap's position is thus untenable: we cannot have philosophy as a theory and at the same time as a fundament to every theory. Either we must give up the theoretical character of philosophy, or we must give up its foundational character. This predicament became central to those philosophers of the analytical tradition who realized the failure of positivism, most significantly to Wittgenstein and Quine. Each of these philosophers chose his own of the two mentioned ways.

Wittgenstein was - in contrast to Carnap and his Vienna Circle colleagues - aware of the predicament from the beginning. He believed in the foundational character of philosophy; and he became clear about having to sacrifice its theoretical character. In *Tractatus* he still believed that philosophy can be a-theoretic, but nevertheless systematic; he considered it would suffice to point out that his system is not a theory in the usual sense of the word, that it is rather a kind of ladder to be kicked away after one climbs it. Later he realized that the notion of such an "a-theoretical system" is doomed, and he started to understand philosophy as something more akin to a collection of "therapeutic" hints.

Unlike Wittgenstein but like Carnap, Quine wanted to have philosophy as a theory. However, unlike Carnap, he clearly saw that in that case he had to give up philosophy as fundamental to science. This led him to reject the very idea of a fundament, the very possibility of Cartesian *prima philosophia*. Hence Quine ends up by considering philosophical theory to be a bundle of scientific results marked not by being firmer or more fundamental than the rest of science, but by being relevant for questions which are traditionally considered as philosophical.

10.8 Logic and Philosophy after the Fall of the Dogma

Our language can be seen as the stage we set up for the world to make its appearance; necessary truth is our setting up of the stage, contingent truths are then the way the world appears.[158] Changing language we change the appearance of the world. However, we must beware of taking this scheme-content way of viewing the language absolutely; it should instead be taken as our way of viewing how language works, as our making sense of the working.

Hence the necessary/contingent boundary is primarily a matter of the outside observer's way of viewing the game; and one and the same game may be viewed in different ways drawing the boundary at various places. This is the point of the holistic insight of Quine and Davidson. However, there is more to necessary truths than this. Some necessary truths are constitutive to the actual language game; and the explicit adherence to them makes it possible for the speakers to retain the common ground. To play the game is to take these truths as necessary. The necessary truth of such statements is not the mere matter of an outside observer's conclusion concerning the way they are handled by the speakers; it is a matter of the speakers' *credo*, of their making it explicit that they are willing to play the language game they play.

The recognition of the real character of the necessity/contingency opposition as initiated by Wittgenstein and Quine, does not have to invalidate logic; but it does involve what many would consider *downgrading* logic. Logic remains a summarization of necessary truths - a finite grasp on the infinity of instances of consequences, a criterial reconstruction of the class of necessary truths; but as necessity turns out to amount not to any ultimate structure of

[158]Hacking (1985, p.155) claims that "although whichever propositions are true may depend on the data, the fact that they are candidates for being true, is a consequence of a historical event." We delimit the space of contingent propositions in that we set up necessary propositions - and in this sense necessary propositions are a purely contingent matter. However, once we let these necessary propositions act as the crash barriers of a language (or of a *style of reasoning*, as Hacking puts it), they acquire their uncontingent, necessary status.

reality, but rather "merely" to the rules of our spelling out of reality, logic turns out to be not the "true canon of the Universe", but rather "only" a code of our theoretizing and argumentation.

What is particularly thrown into doubt is the atomistic view of the world, which has framed our understanding of logic for almost the whole century. Logic still yields cases for the reduction of truth of some statements to the truth of other statements; and these can still be viewed as amounting to the reduction of more complex facts to simpler ones, or to the reduction of some more advanced pieces of our knowledge to other, more primitive ones. However, it seems no longer legitimate to simply assume that there is an absolute, ultimate basis of elementary statements (or of elementary facts, or of elementary pieces of knowledge) to which all other statements (facts, pieces of knowledge) are potentially reducible. What counts as elementary from one visual angle can count as complex from another; and there is no absolute viewpoint, no *God's eye view of the Universe as one closed system* (Putnam, 1984, p.27)[159].

Philosophy can still in a sense be understood as *the pursuit of meaning*; but we cannot understand such a definition as implying that meaning is something absolutely fixed and that the task of philosophy is simply to point it out; and that necessary truth is something given eternally by the way our words hook onto the world and that philosophy can discover the true structure of the world beyond all languages and then the relationship of words to its parts. Philosophy is the matter of a critique of the usefulness of the languages we use, it does not result in theories in our common language, but rather at practical hints that make us see, to use Austin's popular turn of phrase, how we do things with words, and how else we could do them.

We saw that it is not sound to see meanings as pre-linguistic objects which get labelled by expressions; we tried to show that it is more helpful to look at meanings as *values* of the expressions. Expressions acquire values throughout their employment within the process of communication - through being viewed from inside. One of the central tasks of logic, and also of philosophy, is to make these values in some sense explicit.

---

[159]As Hacking (1979, p.315) puts it, "logic, depth grammar, structuralism and the like should postulate points of convergence or condensation, not atoms." For ideas how to dispense with atoms without dispensing with the insights of Russell's and early Wittgenstein's logical atomism, see Skyrms (1993). For a similar point against understanding possible worlds as atoms of the algebra of propositions see Kracht (1993).

11 Language and the World

11.1 Words and Things Pictured

Throughout this whole book we have approached logic as the matter of criterial reconstruction of necessary truth; and we proposed to take meaning to be something that gets explicated precisely in course of such reconstruction. Moreover, in the previous chapter we considered the Wittgensteinian and Quinean challenge to the very concept of necessary truth; and we concluded that it is not legitimate to see necessary truth as reflecting, and hence logic and semantics as explicating, some ultimate "form of the world". Let us now investigate the general picture of the relationship between language and the world, between words and things, to which holding this position leads us.

The commonsense view is that saying language is intimately connected with the world, that it "reflects" the world, means to say that individual expressions reflect individual things - their meanings or their referents. This is to say that the language-world link is usually assumed to be the sum of infinitely many expression-thing links, and this infinite number of links is assumed to be reducible to a finite number of word-thing links. This is the basis of logical atomism as expounded in Section 5.6, and this is what was rejected by philosophers like Wittgenstein and Quine.

If we depict the general language-world relationship as in Figure 4 (amounting to the fact that language reflects the world), then we can see logical atomism as understanding this picture as further analyzable into that of Figure 5 (i.e. assuming that the big arrow of Figure 4 is the resultant of the set of arrows connecting individual things with individual words as pictured on Figure 5).

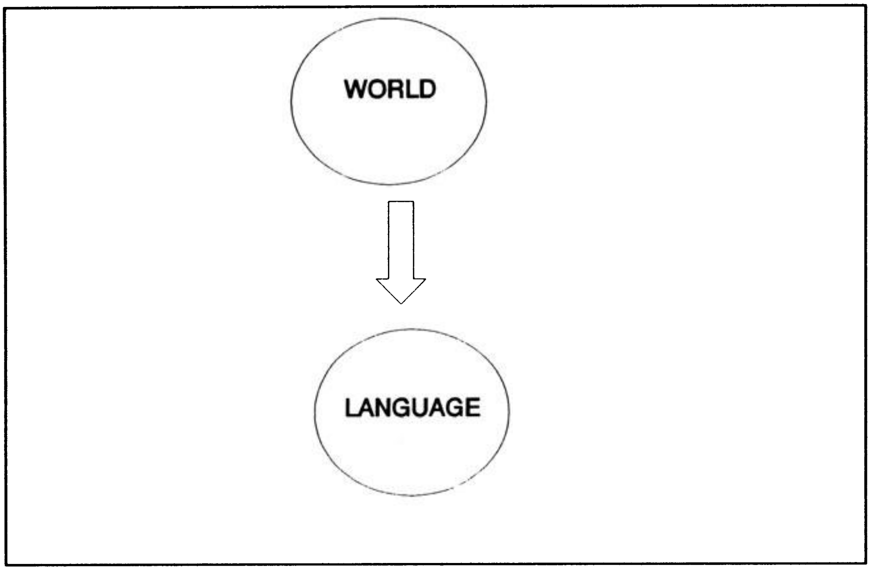

Figure 4

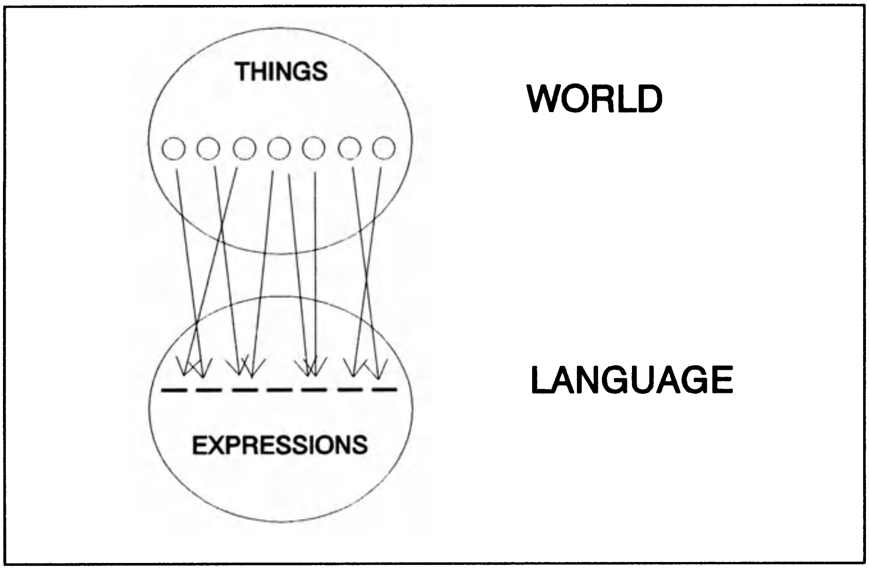

Figure 5

The correspondence between words and things is assumed to be imperfect - different arrows point to the same word (ambiguities) or start at the same thing (synonymies); some words are even without any arrows pointing at them. Thus, language appears to be an unreliable and misleading nomenclature. What seems desirable, and what formal logic appears to offer, is a better, exact nomenclature, providing one and only one word for every relevant thing. So the role of formal logic is taken to be to *replace* natural language (at least where scientific aims are pursued); as in Figure 6[160].

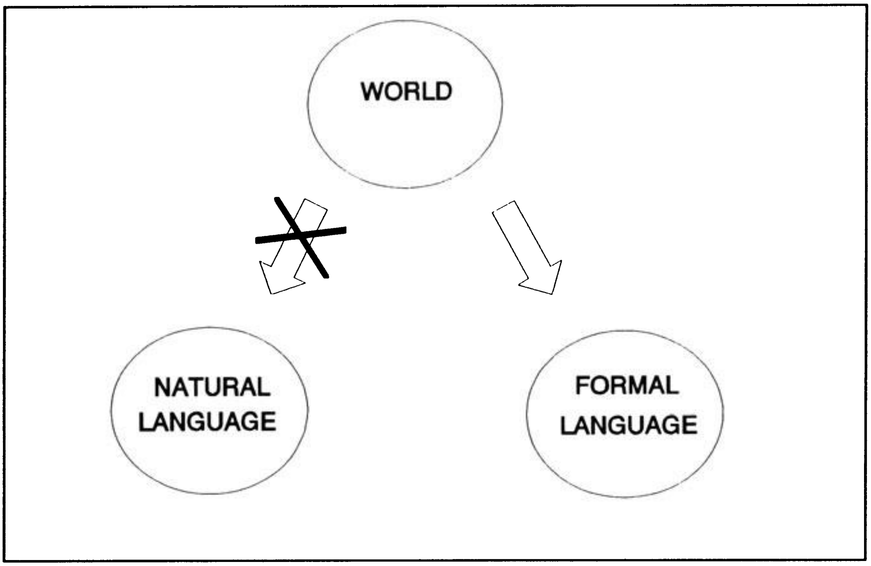

Figure 6

Model theory seems to bring a further refinement to the process of making our nomenclature more precise - it seems to allow us to first reconstruct the world by means of model structures, which appear to allow for a kind of direct materializing of our imagery, and then to devise the language suitable for treating of the model structures. Figure 6 thus gets refined to Figure 7.

[160]This is, of course, not to say that logicians and semanticists expect people or scientists to abandon their ways of talking and to start talking some language of formal logic or set theory instead. It is to say that they take the translation of people's talk into their logical notation as ultimately revealing what the talk is *really* about.

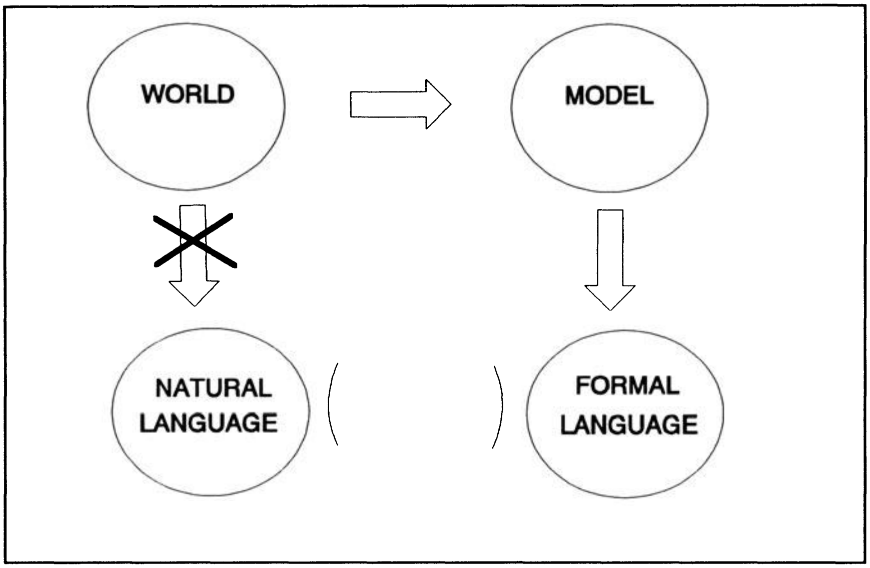

Figure 7

We use set theory to make the precise model of the denizens of the world (plus, as the case may be, of all the worlds we can think of), and we use this model to explicate the semantics of formal logic and to purify its expressive means in order to reach the right nomenclature. Formal logic thus purified can then be used to analyze and to rectify natural language. Hence according to this view, logic is not a regimentation of natural language, it is a (preferable) alternative to natural language. It is superior to natural language and it is thus the norm in an absolute sense.

It was this picture we have tried to challenge in the preceding chapters. Instead of embracing logical atomism we have pointed out the weight of Wittgenstein's, Quine's and Davidson's holism. We tried to show that formal logic can be seen as the fruit of our efforts to envisage the "inferential potential" of our language, and not a means of rendering some "true structure" of the world. We tried to point out also that model theory should not be seen as a tool for doing metaphysics (or Wittgenstein's ultra-physics), but as only a sophisticated means of "visualizing" those aspects of language which are relevant from the point of view of inference. Hence we proposed that the relationship between natural and formal languages should be seen not as that on Figure 7, but rather as that on Figure 8.

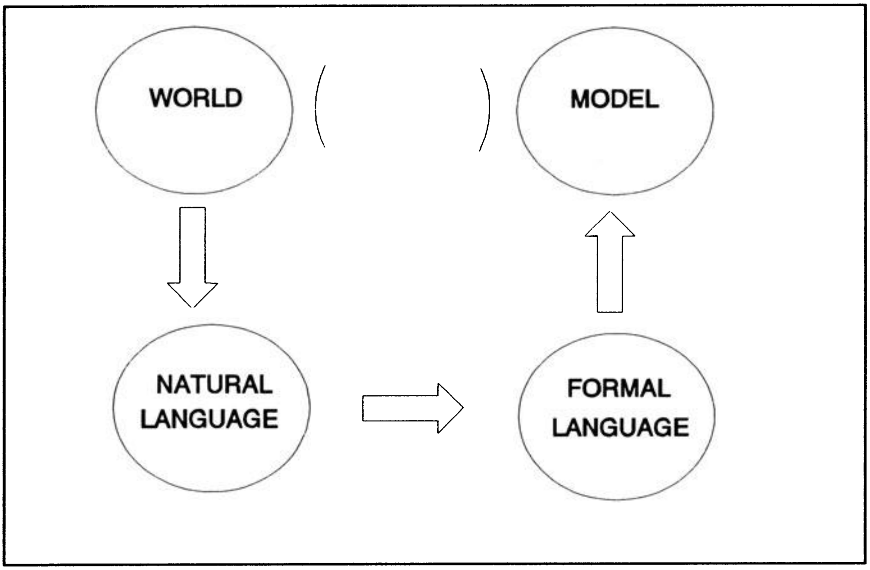

Figure 8

Logic arises out of regimentation of language; and model theory arises out of further elaboration of logic. Logic is language stripped of everything that is irrelevant from the point of view of inference; model theory then rearranges the inferential potential so that it gets objectualized, and hence yields the 'natural language metaphysics'. Model theory does - in a sense - reflect *what there is*, not however by way of mirroring or of direct modelling, but rather by a distillation of linguistic articulation. "The suggestion is," as Davidson (1977b, p.201) puts it, "that if the truth conditions of sentences are placed in the context of a comprehensive theory, the linguistic structure that emerges will reflect large features of reality."

9.1 Philosophy and The Linguistic Turn

To increase our insight into the real nature of language and the role of logic, it may be helpful to look at the sources of the crucial changes undergone by philosophy and logic in this century and to see how it came that many representatives of these two enterprises found them converging to intersect in the field of what is usually called logical analysis of language.

Language has been traditionally assumed to be a tool peculiar to human race. Just as our prehistorical ancestors discovered that a task they were to perform, such as hunting, can be facilitated by employing certain suitable tools, such as spears, axes or bows, we imagine our perhaps a bit more recent prehistorical ancestors as discovering that another task, namely communicating ("letting others know the content of one's mind"), can be facilitated by employing another tool, namely language. Language has been thought to be the means of labelling the contents of one's mind or entities of the outer world; and presenting the world "as it really is". This led directly to the nomenclatural view of language.

This commonsense view of language has been always hailed by some philosophers, and questioned by others. Our century can be thought as outstanding in bringing new ways of challenging it. There has been a particularly wide change of mind - the linguistic turn - affecting a broad range of philosophers, and there was a vigorous development of the entirely novel and inspiring way of studying language based on formalization; and these impulses opened quite new vistas on language and on its relation to the world. However, the enthusiasm brought about by this vigorous development also brought about much confusion about what a theory of language can and should achieve.

The enigma of the language-world relationship was brought to the centre of philosophical discussion early in this century by Frege, Russell, Wittgenstein, Austin and others. Their original point was that we cannot take the representing capacities of language at face value, that in order to treat of things - which cannot be done save with the help of words - we must first treat of words and make sure which of them are really capable of treating of things. This is what the term *linguistic turn* was introduced to express: philosophers slowly stopping asking *what is consciousness (matter, evil, etc.)?* and instead beginning asking *what is the meaning of 'consciousness' ('matter', 'evil', etc.)?*

Those who remained with the nomenclatural notion of language interpreted the linguistic turn as disclosing the deceptivity of the language-nomenclature and as pointing out the necessity of a better, "exact" language. Russell, Carnap and other philosophers concluded that the problem with language is that expressions which appear to stand for an object may well not stand for any. In his path-breaking paper *On Denoting*, Russell (1905) used "logical analysis" to point out that despite appearances, expressions such as *someone*, *everyone* or *the king of France* should not be considered as names; Carnap (1931) then used similar means to substantiate his claim that such "metaphysical words" as *principle*, *God*, *idea*, etc. are in fact without any meaning and that they act only as a certain kind of externalization of the speaker's feelings. Such train of thought led to the view that all there is for philosophy to do is to supply a new universal nomenclature which would be better than natural language; a nomenclature free from the flaws which had given rise to traditional philosophical (pseudo)problems - and it was assumed that it is just this that modern formal logic is (successfully) targeting. This was the point where many philosophers started to perceive logic as swallowing up philosophy in the traditional sense; and where they became "analytic".

11.3 Logic and the Formalistic Turn

The reason for the sudden increase of logic's significance in philosophy was not only due to the linguistic turn having established the need for a "better" language, but also because logic itself was undergoing a change which particularly suited it to serve this need. This change can be called the *formalistic turn*.

In Section 2.7 we stressed the importance of distinguishing between using logical symbolism as means of mere regimentation and schematization, and using it as a means of formalization. We can speak about *symbolic* logic in the first case, and about *formal* logic in the second. Doing symbolic logic we use symbols to regiment labyrinthine natural language, to suppress that which is irrelevant from the point of view of consequence and to emphasise that which is relevant. We use them also to articulate types of language expressions to summarize patterns of consequence. Doing formal logic, on the other hand, we liberate logical symbols from the bondage of their natural language prototypes and grant them life of their own. We take systems of logical symbols - formal calculi - as self-contained algebraic structures and we may use them for normative purposes.

We have seen that Frege's original intention was to do symbolic logic; and the same can also be said about Russell. For them, a logical formula was simply a regimentation of a pre-logical statement and hence amounted to a definite proposition or a definite "thought".[161] "Jeder, der Worte oder mathematische Zeichen gebraucht," Frege (1884, p.22) writes, "macht den Anspruch, daß sie etwas bedeuten, und niemand wird erwarten, daß aus leeren Zeichen etwas Sinnvolles hervorgehe."[162] There was little sense in considering varying truth-valuations and reference- or denotation-assignments.[163] The alternative approach to logical formulas, based on viewing them as abstract objects capable of being interpreted in various ways, was - to a certain extent - implicit in the approach of the "algebraic logicians" such as Boole and Schröder; but it was Hilbert who first applied it systematically and who thereby started to do logic that was really formal.[164] This move was a natural one: logical theory had reached a degree of complexity which demanded the separation of questions concerning the apparatus from questions concerning its application, i.e. the separation of the inner problems of logical cacluli from the problems the calculi were originally devised to elucidate. However, formal logicians soon came to be so preoccupied with solving problems of the former kind that many of them began to overlook that the genuine problems for logic to solve are those of the latter.

As soon as formal calculi came to be considered as self-contained mathematical structures, they started to reveal a wide range of mathematically interesting problems. One class of problems concerns the character of the class of theorems: Is there a statement which would be contained in the class together with its own negation, i.e. is the calculus consistent? Does the class contain the negation of every statement not contained in it, i.e. is the calculus syntactically complete? What is the cardinality of the class; what are its structural properties?

---

[161]For both Frege and Russell, symbols were, as Tichý (1988, p.ii) puts it, "not the subject matter of their theorizing but a mere shorthand facilitating discussion of extra-linguistic entities." Tichý himself is probably the most consequential contemporary continuator and the most diehard advocate of the Frege-Russell tradition in logic.

[162]"Everyone who uses words or mathematical symbols makes the claim that they mean something, and no one will expect any sense to emerge from empty symbols."

[163]"For Frege it cannot be a question of changing universes. One could not even say that he restricts himself to *one* universe. His universe is *the* universe." (Heijenoort, 1967, p.325)

[164]The nature of the difference between Fregean symbolic and Hilbertian formal logic is clearly seen when we consider the controversy between the two logicians about the nature of axioms as expounded in Section 3.4.

Another range of problems concerned the way the class is specified, i.e. the axiomatic system: Are the present axioms and rules the "simplest" way to determine the theorems, in particular would it not be possible to omit an axiom, i.e. is the system independent? Would it not be possible to omit an inference rule? However, the most intensively discussed range of problems concerned the relation of calculi to their set-theoretical interpretations as studied by Löwenheim, Skolem, Gödel, Tarski and their followers: Are all theorems verified in all possible domains, i.e. is the calculus semantically sound? Are, vice versa, all statements verified in all possible domains theorems, i.e. is the calculus semantically complete? Is there a domain that would verify all and only theorems?

The formalistic turn meant the step from symbolic to formal logic, from van Heijenoort's notion of *logic as a language* to his notion of *logic as a calculus*[165]. It completely changed the character of logic - what was originally a mere tool of logic was slowly turned into the subject matter of logic. Moreover, model theory, surely the most spreading and the most fascinating outgrowth of the turn, presented the next step in the takeover of philosophy by logic: after the logical analysis of language as pursued by Frege, Russell and Carnap eliminated the old metaphysics, Tarskian model theory provided for a new metaphysics, now in a mathematical disguise. This trend caused many logicians and philosophers simply to identify philosophy with logic[166]. They came to believe that traditional philosophical problems emerged only because philosophers of the past did not have the apparatus of modern formal logic.

11.4 Capturing Meaning and Logical Analysis

Elsewhere (Peregrin, forthcoming c) I put forward the claim that whereas ordinary communication, the "normal" usage of language, yields grasping the relation between an expression and its meaning as an immediate, "a priori"

[165]See Section 2.7.

[166]The identification taking various shapes, from Carnapian "philosophy as the logic of science" to Montagovian identification of model-theoretical semantics with "formal philosophy".

matter, the common denominator of the linguistic and the formalistic turn consists in suppressing this stance and viewing the relation instead as a fact which is contingent, "a posteriori". To view our language in the latter way means to view it as if it were a foreign language, a language we do not understand; to perceive its expressions - contrary to the usual way - as not inherently equipped with their meanings.[167] This institutes an ambivalent approach to language - our language, which necessarily keeps acting as a medium of shaping of all the things of our world, comes to be also considered as one of the many things within the world. We may accept Tarski's advice and distinguish between the object language (the thing) and the metalanguage (the medium); however, we have seen that to treat natural language as the object language necessarily invovles gerrymandering it in such a way that it is deprived precisely of what makes it natural.

Reassuming the spatial metaphor introduced in the previous chapter, we can say that whereas ordinary communication means being inside language, the linguistic turn requires us to look at language from outside. The same is required by the formalistic turn in respect to the language of logic - we have to cut off logical formulas from their natural language counterparts which they were originally devised to regiment. However, investigating language from outside, taking it as an object, is simply doing empirical linguistics. This is not what could underlie a revolution in philosophy envisaged by the linguistic turn. To make sense of the turn we cannot simply suspend taking language as a medium, we need it precisely in its role of the universal means of grasping the world. The essential error of some of the followers of the turn was failing to realize this - they came to believe that the consequence of the linguistic turn is that all that is required for solving traditional philosophical problems and for getting a grip on meaning, is to do linguistic analyses or to play with formal calculi.

So it seems that what we need is to capture meaning without going outside of language, without letting language degenerate into one of the vast objects of the world. Now let us realize that this is precisely what *symbolic* logic could help us do. Let us notice that it is not the Fregean symbolic approach to logic, but really only the Hilbertian formalistic turn that demands the "view from outside". We have seen that the primary aim of logic is to

---

[167]Using Goodman's (1984) terms, we can say that it means thinking *of* words without thinking *in* the words.

summarize basic instances of consequence, basic patterns of our reasoning used in arguments and proofs. The use of symbolic devices within logic arises from recognition of the fact that the patterns are easier to summarize if we do not take natural language at face value, but reconstruct it instead as a strictly rule-based grammatical system. It does not require us to go outside language; what we do is to replace natural language statements and arguments by their formal regimentation which allow for ignoring irrelevant idiosyncrasies and seeing the relevant regularities. Logical schematization understood in this sense means *capturing the unity of meaning within the multiplicity of surface forms* via accounting for the infinite class of valid instances of consequence by finite means; and logical analysis amounts to collating language with this reconstruction.

However, once interpreted formal calculi come to be perceived as alternatives to natural language rather than as mere means of its regimentation and reconstruction, logical analysis comes to be seen not as a matter of the comparison of language with its meaning-capturing, criterial reconstruction, but rather as a way of making the meanings of expressions explicit by furnishing them with model-theoretical interpretations. The problem of explicating meaning thus comes to be understood as the problem of finding a model theory adequate for natural language, as a "representational semantics"; and this *does* mean moving outside language. We must, it is claimed, first develop adequate set-theoretic representations of what the world is like and what it could be like, and only then study the relations of statements to these representations. We seem to be able to go through our "mental images" of how things could be and to model these images by means of model structures.

To explicate a phrase of a foreign language it is enough to display the corresponding phrase of our own language; and it may seem that to explicate a phrase of our own language it is likewise enough to display the "direct", model-theoretical capturing of the real-world situation, or the mental image, which this phrase appears to represent.

11.5 The Ineffability of Semantics

We may say - with Tarski - that *Aristotle is human* is true if and only if Aristotle is human. This is, of course, true, but equally of course, it is completely trivial, it merely points out the disquotational character of the predicate *is true*. We saw this in Chapter 6.

However, it might seem possible to insist that such a Tarskian T-sentence means something over and above this, that it says that if *Aristotle is human* is true, then it is *due to the fact* that "out there in the world" or possibly in a model structure which is considered to offer a faithful representation of the world there is an individual Aristotle and this individual instantiates the property of being human, or that Aristotle is an element of the set of human entities, or that there exists a fact of the coincidence of Aristotle and humanity. However, saying this, saying that the statement is true iff, e.g., the individual Aristotle instantiates the property of being human, explains nothing; it merely offers a cumbersome paraphrase of the statement[168]. Such a paraphrase might have a purpose, namely comparing language with some reconstruction, in this case with a model-theoretical reconstruction; but in such a case it makes sense only globally, not locally. It can give neither the ground of the truth of a single statement, nor its meaning. The general reason is that if we put aside translation, then we are simply not able to give the meaning of an individual expression or truth conditions of an individual statement in a nontrivial way.

How do we learn the meaning of an expression? There are two principal ways: we can be shown how to use the expression (in some particular cases which kinds of things this expression applies to), or we can be provided with the translation of the expression into a language with which we are familiar. It is important to realize that the first of these ways is an essentially *practical* way; and it cannot be fitted into a *theory*. I can show someone what *human* means, but if I want to put it down, I must somehow express what it is that I was showing and I end up with either the tautological "*human* means human", or with the less trivial "*human* means such-and-such", which gives the meaning of one term by translating it into another term and which thus cannot be used generally - on pain of the vicious circle.

---

[168]And accepting it as an explanation is, as Rorty (1989, p. 8) puts it, like explaining why opium makes you sleepy by talking about its dormitive power.

This indicates that what logical analysis can bring to light are always *relations* between expressions (especially the relation of consequence between statements) and consequently the structure of language, not meanings of individual expressions. Equating an expression with a formula of a formal calculus is not to pick up the thing the expression stands for, but rather only to point out the place of the expression within a certain system which the analysis identifies on language and which the calculus materializes.

Let us illustrate this using possible worlds semantics. Within such semantics, the meaning of *Aristotle is human* would be the set of those possible worlds in which Aristotle is human, i.e. {w| Aristotle is human in w}. But to understand which set is meant we would have to understand the expression *{w| Aristotle is human in w}*, and in particular we would have to understand *Aristotle is human*; so the theory would give us the meaning of *Aristotle is human* only if we knew the meaning in advance. Someone might object that the class of relevant possible worlds should be specified in a more elementary way, that the specification should be {w| Aristotle is such-and-such in w}, where *to be such-and-such* stands for some explication of what it is to be human; but in such case what is relevant from the point of view of the meaning of *Aristotle is human* is the purported synonymy relation between *Aristotle is human* and *Aristotle is such-and-such*, not the possible worlds. Thus, possible worlds do an interesting work only in the context of language as a whole (or a nontrivial part of it); again, they help to bring to light certain structural properties of language, not meanings of individual expressions. And what holds for possible worlds, holds for more sophisticated tools of formal semantics as well.

The nomenclatural view of language may be understood to imply that it is the task of the linguist to take an expression and put it into his theory side by side with its meaning.[169] But this is, as we have just seen, simply impossible - the only thing that can be the subject matter of linguistic theory are - ultimately - relations between expressions, not relations of expressions to things. This is not to deny that it sometimes is useful to view language as a nomenclature; many aspects of language, relevant especially for the methodology of science, can be made intelligible in this way. But there are other aspects of language which escape such an approach, and if what we are after is a fully-fledged theory of language, then we may make use of the nomenclatural view, but we cannot be held captive by it. Sometimes it is useful to view language as a

[169] See the suggestion of Tichý (1992) and my criticism (Peregrin, 1993c).

nomenclature; but on other occasions it is much more useful to view it, following Wittgenstein, as a kind of toolbox.

If we view language as a peculiar kind of human behaviour and the theory of language as a theoretical reconstruction of this behaviour, i.e. as the enterprise of finding rules which would "account for" the behaviour (i.e. which would generate the same patterns of behaviour as those factually observable), then we reach a picture essentially different from the nomenclatural one. We do not assume that language has its origin in baptizing things - an obscure assumption, of which the possibility of justification is notably unclear. Certainly, we may base the rules of our theoretical reconstruction of language on relating expressions to some entities, but there is no sense in considering this move as "the discovery of the true nature of language", let alone that of the world; we must consider it merely as one possible approach to language yielded by one possible point of view.

If we realize that our language is the "universal" (the illuminating German word *unhintergehbar* unfortunately has no exact English equivalent) medium, then we must conclude that semantics is essentially ineffable - to be able to get an explicit grasp of meaning we would have to step outside language, but making this step means downgrading language to an insipid thing. This is to say that we cannot go outside, that there is in fact no outside. "Es gibt gar kein draußen;" as Wittgenstein (1953, §103) puts it, "draußen fehlt die Lebensluft."[170]

By providing a model-theoretical interpretation for a formal calculus or for a natural language we offer a new perspective which may help us perceive patterns and regularities which would remain hidden to our eyes otherwise; however, it is inadequate to see this act as the act of going from the words to what the words are about.

---

[170] "There is no outside; outside you cannot breathe."

11.6 How We Make Worlds with Words

We can indeed make worlds with words; moreover, in a sense there is no other way of making worlds. Our ways of "worldmaking", as expounded by Goodman (1978), are essentially involved with words.

However, it is not only that a concrete world has to be fabricated with the help of language; language is also that which is necessary to open the realm of possibilities which can then be captured in terms of the space of possible worlds. As Wittgenstein pointed out and as we saw in Section 10.5, what is possible is inseparably connected with that which we accept as expressible. A space of possible worlds is always only one side of a coin the other side of which is a language. To conceive of an absolute space of possible worlds means to daydream about being outside all languages; and it inevitably means to get trapped into antinomies.

Let us, for the sake of illustration, consider Putnam's (1985) point about the Tarskian notion of truth. Putnam claims that if Tarskian T-sentences for a language L are theorems of logic in meta-L, then we must conclude that these sentences hold in all possible worlds. This means that *'Snow is white' is true if and only if snow is white* holds true in all possible worlds, including, Putnam claims, those worlds in which the statement *Snow is white* does not mean that snow is white.

We have seen that the status of the T-sentence *'Snow is white' is true if and only if snow is white* depends on our position with respect to the language from which *Snow is white* comes. If we are inside this language, then the T-sentence is necessary; but then we also do not see a world in which the statement *Snow is white* could mean something other than that snow is white. We simply cannot distinguish between the statement and its meaning (we *grasp* the statement inseparably together with its meaning). On the other hand, if we are outside the language, then we can see worlds in which *Snow is white* means even the most strange things; then, however we also do not grasp the T-sentence as necessarily true.

This indicates that the space of possible worlds is always language-relative; there is always a language which is "about" all the possible worlds, but which is not inside these possible worlds.[171] Let us, for the sake of argument, suppose that we have an absolute, all-embracing space of possible worlds. Then, for every statement of our language there are possible worlds in which the statement means other things than what it means within our actual world, in particular there are worlds in which it means something false. Hence there are no necessarily true statements, and there is nothing for logic to study. Moreover, as meaning is inseparably connected with necessary truth, there is no meaning and hence no theory of meaning. Thus, the allembracing space is a useless, trivializing illusion, an illusion of the Archimedean point (being outside of every language), which Putnam (1988, p.89) himself explicitly rejects[172].

A space of possible worlds is only a general way of an unfolding of the inferential properties of a language - it is a general way of reification of necessity as expressed in the language, just as the universe of discourse is the unfolding of certain more specific inferential properties of language. We are often inclined to regard possible worlds as something at which we can directly look in the same way as we look into various rooms of a house, and this makes us believe that the possible worlds can be employed to give necessary truth an independent check. However, the independence is an illusion. *In every possible world* is nothing over and above *necessarily*[173].

In the previous chapter we concluded that necessity can be seen as contingency viewed from inside. Now we can add that it is language that makes such a "view from inside" possible in the first place. We can think of language

---

[171]Cf. Mates' (1968) distinction between being true *of* a possible world and being true *in* the possible world.

[172]The fact that we can speak about possible worlds means that possible worlds exist - in a sense - within our (actual) world. This may mislead us into seeing necessity as a peculiar kind of contingency. This line of thought leads to what McGee (1992, p.274) calls, inspired by Etchemendy (1990), the *contingency problem*: to be valid is to be true in all models; but which models exist is the matter of contingent fact. However, to inquire whether necessities are not after all in some sense contingent is essentially erroneous - necessity is what makes every contingency possible in the first place. To examine necessary truths within the framework of which they are constitutive is like asking whether logic is logical, or, to use Wittgenstein's (1953, §50) well known example, whether the meter-standard is one meter long.

[173]As Kripke (1972, p. 267) it: "'Possible worlds' are *stipulated*, not *discovered* by powerful telescopes.

as of something which is characterized by this ability, in contrast to other entities, to create this kind of inside, an inside - to put it poetically - to harbour human selves.

We can imagine that seen "from the Archimedean point", "from Nowhere", or "by a God's eye view", the world can be considered as sheer contingency displaying regularity of only the causal kind; linguistic utterances appearing to be merely peculiar kinds of events obeying the all-encompassing web of causes and effects. But insofar as we are no Gods, each of us has to dwell "within a language", to observe the world through its prism and to perceive some of its God's-eye-view-contingencies as necessities. We may hold that our human perspective is narrow and parochial, and we may construe Quine as saying that there are "in fact" (which can mean nothing else than: seen "by a God's eye view") no necessities; but I vote for discarding this unattainable vantage point and for reconciling ourselves to our human deal with its limited perspectives and with its parochial necessities.

11.7 The Right Language?

Our crucial claim is that model theory should be conceived of not as a way of modelling of the language-independent world, but rather as the result of a certain way of regimentation of language; that it is not adequate to consider model theory as reflecting, and logical analysis as collating language with, "the structure of the world". The reason is that our grasping of the world is always mediated by a language. However, it may be objected that our point is a red herring, that it makes no real difference whether we talk about the structure of the world or the structure of language.

The objection might run as follows: As language is bound to reflect - however distortedly - the world, its regimentation, if carried out in the "right way", must lead to the "right" structure, to the same structure which would result from the direct nomenclatural capturing things. And so even if you deny the possibility of a direct access to the world and maintain that the only thing we can do is to analyze language, there seems to be no reason not to call the structure which you reach by means of every "right-headed" analysis of language (i.e. the one which has been, or is to be, discovered by the "right" logic) the structure of the world. From this perspective, the difference between

Figures 8 and 9 is not substantial - it is the difference between two possible methodologies, which are bound to converge on the same result.

The trouble with this objection is that it presupposes that there is something that could be reasonably considered as *the* structure of language, that there is one and only one "right" way to "represent the world". It assumes that the general structure of every natural language, which is brought to light if we purify it of local distortions and misrepresentations, is bound to be one and the same. It is true that if there were something which could be really called *the* structure of language, then there would indeed be no reason not to see it as the structure of the world; however, to assume such a unique structure is rather visionary. For one thing, it is not clear why every foreign language would have to be essentially like ours; and, moreover, every structure is always the result of an analysis, and it is not clear why there should be one and only one right way to analyze language.

What may be the reason to believe that there is one and only one right way to "represent the world", one and only one "true" structure of language? Well, we may somehow believe that the world itself is a kind of language and that our human languages are mere better or worse approximations of this *mathesis mundi* (a more - or less? - far-fetched variant of this belief is the belief that the world is not itself a language, but that there is a language that is forced upon us by the world[174] or by its God). In fact, it is an idea very much like this that lies the foundations of logical atomism. However, although this might have great appeal for a mystic, it is hard to see reasons for a rational philosopher to embrace it.

A non-mystical alternative to this idea is the idea of "the True Language of Science", the idea that our science is able to decide how the best language with which the world could be described should look like. This is the idea put forward, e.g., by logical positivists during the first half of this century; but the positivists clearly failed to give it a substantiation. In fact, instead of pointing out the right language to describe the world, the positivists succeeded only in making it clear that there is no such single language; in particular that there is no objective ground to make the decision between the physicalistic and the phenomenalistic language. And it is hard to see the rest of this century as bringing any change to this - on the contrary, the well-known puzzles of subatomic physics indicate that the idea of a Single True Language is even

[174]Goodman (1984, p. 21) speaks about "inexorable dictates of nature".

much more far-fetched than it appeared after the failure of logical positivism. And to see this we need not subscribe to such extreme sceptic views of science as those put forward by such philosophers as Foucault (1966) or Feyerabend (1978).

Moreover, the results of comparative linguistics as presented by Sapir (1921), Whorf (1956) and their followers appear to bring lots of empirical evidence against the conviction that all natural languages are essentially alike. They put into doubt the assumption that all languages share a common structure, let alone that their common structure would be that of traditional logic.[175] And again, even if we do not embrace the Whorfian radical scepticism about the commensurability of structures of individual natural languages, we can hardly disregard such facts as that there are languages which lack anything similar to determiners like *every* or *some* basic to English and other Indo-european languages and constitutive to our present logic.[176]

Maybe this is still insufficient to prove a linguistic relativism - for to give an analysis of an unknown language is an intricate business and it can be argued that Whorf and others who found some languages incommensurate with English did so only because they failed to discover the right way to analyze them. However, even if we consider Whorfian arguments as not persuasive enough to substantiate linguistic relativism, we simply cannot take the contrary for granted. This is to say that even if we do not consider arguments in support of incommensurability of different languages compelling enough, we can hardly take commensurability as anything more than an empirical hypothesis which we expect not to get disconfirmed; we cannot make it into a pillar of our view of the world.

And if we give up the axiom of one right language, then we are left with that way of making sense of logic which we proposed above. If we see the language of logic which we use to do semantic analysis as better than the analyzed natural language, then it is not because it would be "closer to the

[175]Whorf (1956, p.213) claims: "We cut nature up, organize it into concepts, and ascribe significances as we do, largely because we are parties to an agreement to organize it in this way - an agreement that holds throughout our speech community and is codified in the patterns of our language."

[176]This fact is, I think, quite clear to linguists grown within the European tradition (see, e.g. Hagège, 1993); but now it comes to be appreciated also by many American, even formally-minded, linguists (see, e.g., Bach et al., to appear.).

world", that it would be a more precise nomenclature, but because it is in some sense *canonical*. What this means is that it provides us with some kind of schematic rendering of the whole of language, while it is considerably simpler and more comprehensible than the language itself and nevertheless retains its semantically relevant properties. Doing this helps us understand language, because it facilitates comprehensibility of a certain *structure* of language (understanding structure, of course, more inclusively than syntactic structure) and of the way language functions.

Thus, we are back to what we called structuralism: to explicate language is to carry out its criterial reconstruction, to visualize a structure; to explicate the meaning of an expression then is to point out the position of the expression within the structure thus visualized. If the reconstruction is carried out by way of logical regimentation, then pointing out the position simply means finding the particular formula which regiments it. To explicate the statement *Aristotle is mortal* means to point out the formula **Mo(Ar)**, but not because this formula would point out the piece of the world corresponding to the statement in a more direct way than the statement itself, but because it associates the statement with the position within the structure envisaged by the predicate calculus, i.e. in the structure capturing the basic "inferential potential" of language.

11.8 Correspondence Trivialized

There is of course no quarrel that there is a sense in which language does "reflect" the world. The naive view is that language is a copy of the world, that we make language to be able to treat of things *in absentia*, in the same way as we make a map to be able to treat of a land without having to travel to it. This idea leads to perceiving language in terms of representation, and thus it lends credibility to the nomenclatural notion of language.

However, we have seen that this notion is dubious, as the structure of the world is not definite in a way which would allow us simply to pick up its elements and to pair them with language expressions. The reason is that the very structure of the world, and hence what counts as an element of the world, is always biased towards the language we use to talk about the world; and there is no absolute, neutral and unbiased world. This precludes the possibility of considering correspondence between language and the world as something substantial. "There is no such thing," as Goodman (1960, p.56) emphatically puts it," as the structure of the world for anything to conform or fail to conform to."

The correspondence theory of language and of truth (or the *picture theory* as it is sometimes called in connection with the early Wittgenstein; or the *copy theory* as Putnam, 1981, calls it) draws on the intuitive feeling that statements are pictures of facts in very much the same way in which drawings and paintings are pictures of real things and scenes.[177] However, let us notice the fact which is quite banal, but which has got completely obscured by the picture-metaphor: a statement does not resemble a fact in any way really analogous to the way a drawing does. A statement and a fact are two entities of such entirely different kinds that the former can - save perhaps in some very special cases - not literally *picture* the latter (in the sense of being its *icon*).

However, this obvious fact is often commonly denied, and statements are taken to really picture facts. Constituents of a statement, i.e. words, it is said, are names of constituents of facts, i.e. of things, and the syntactic structure of a statement coincides with the structure of a fact. The statement *Aristotle is mortal*, for example, reflects the fact that Aristotle is mortal: *Aristotle* is the name of Aristotle, *is* expresses something like coincidence or instantiation, *mortal* is the name of mortality; and the fact that the three words are connected into the statement pictures the fact that the three entities which are expressed by them are connected in the corresponding way into a fact, namely into the fact that Aristotle is instantiating mortality.

This notion amounts to the following picture: we *perceive* how the world is, and in particular we perceive various entities and the ways they are configured. We have employed words to baptize the entities and now we use the conglomerates of words which correspond to the conglomerates of things perceived, thus externalizing our perceptions. However, this picture is wrong, and it is wrong so obviously that once we see it, it is hard to understand why it has come to be so often taken for granted: we do not *perceive* Aristotle, being and mortality as linked one to the others; neither do we perceive Aristotle and mortality as being linked one to the other by means of a "being". We can in no sense say that we perceive a situation consisting of three parts corresponding to *Aristotle*, *is* and *mortal*.

---

[177]This is indeed the analogy that gave birth to the correspondence theory of Wittgenstein's *Tractatus*.

The perception of the mortal individual Aristotle is in no such way structured, it gets structured only in so far as it becomes classified by a linguistic utterance. Moreover, few statements can actually be thought about as intimately related to a perception in this way; and to model language on "externalizing perceptions" is thus misguided (Wittgenstein devoted much of his late work to showing this).

In his well-known critique of empiricist epistemology and of what we call the nomenclatural notion of language, Sellars (1956) pointed out that this notion cannot capture how language really works, because between a perception, or any kind of a sense datum, and an observational *knowledge* there lies a gap; an observational knowledge is neither identical with, nor trivially yielded by, a perception, it is something which arises from the perception only by way of a nontrivial engagement of the "classificatory powers" of language as a whole. As Sellars (ibid., p.176) puts it, "instead of coming to have a concept of something because we have noticed that sort of thing, to have the ability to notice a sort of thing is already to have the concept of that sort of thing, and cannot account for it".[178]

Therefore, it is not feasible to see language as repeating an intrinsic structure of the world, it is more adequate to see it as *giving* the world its structure. The problem is that we live in the world which is structured by our language all the time, we *see* the world as having this structure and it is hard for us to imagine that this could not be *the* structure of the world. We assume that this structure is intrinsic to the world and that our language is devised to fit it. Sellars (1956, pp.161-162) warns us of this delusion; he points out that "unless we are careful, we can easily take for granted that the process of teaching a child to use a language is that of teaching it to discriminate elements within a logical space of particulars, universals, facts etc., of which it is already undiscriminatingly aware, and to associate these discriminated elements with verbal symbols."

We look at a statement and then at a part of the world to which the statement seems to correspond - and we see the two items as sharing a common

[178]The problems connected with believing in "purely observational knowledge" in the context of scientific discovery were discussed in detail by Hanson (1961). It is interesting to note that the disbelief into the positivistic notion of observation was expressed already by Frege (1884; p.99 footnote): "Das Beobachten schließt selbst schon eine logische Tätigkeit ein." ["Observation itself already includes within it a logical activity."]

structure. However, we might not notice that the structure of the latter is a mere imprint of that of the former and that the coincidence is thus trivial. The problem is that we simply cannot help seeing the world structured in the way we see it, and once we begin to look at our language "from outside", we are purporting to see *two* kinds of structure: a structure of the world and a structure of the language. We find these two kinds of structures alike and we begin to speak about correspondence. However, there are no such two kinds of structures, there is only a twofold perception of one and the same structure. Even if we look at our language "from outside", we in a sense still remain within it, and our language continues to be what organizes our comprehension of our world. We cannot simply go out of our language; we are held captive by it. This is what is pointed out by Wittgenstein (1953, §115):

> Ein *Bild* hielt uns gefangen. Und heraus konnten wir nicht, denn es lag in unsrer Sprache, und sie schien es uns nur unerbittlich wiederholen.[179]

Moreover, as we pointed out, no structuring of language, which gets so imprinted onto the world, need be shared by every conceivable language; different languages need not be alike enough to let all people perceive the world in the same way.[180]

---

[179] A *picture* held us captive. And we could not get outside, for it lay in our language and language seemed to repeat it to us inexorably.

[180] This has been urged by Whorf (1956, p.221): "The phenomena of language are background phenomena, of which the talkers are unaware, or, at the most, very dimly aware - as they are of the motes of dust in the air of a room, though the linguistic phenomena govern the talkers more as gravitation than as dust would. These automatic, involuntary patterns of language are not the same for all men, but are specific for each language and constitute the formalized side of the language, or its 'grammar' - a term that includes much more than the grammar we learned." Putnam's (1990, p.28) philosophical analysis leads to a similar conclusion: "Elements of what we call 'language' or 'mind' *penetrate so deeply into what we call 'reality' that the very project of representing ourselves as being 'mappers' of something 'language-independent' is fatally compromised from the start.*"

11.9 Formal Logic as 'perspicuous representation'

Thus, the purpose of formalization is to help us see certain aspects of language, namely its "inferential potential", and its functioning more clearly, to provide for what might be called *perspicuous representation.*[181] It is like a prism through which we perceive the factual language to help us see the regularities of its functioning.

This brings about a concept of logical analysis quite different from that put forward by Russell, Carnap and the other pioneers of analytical philosophy. These philosophers were convinced that the structure of language is something quite definite, something duplicating the structure of the world, and that the task of logical analysis is to dig this regular structure out from beneath the irregular surface of language where it lies buried. There was assumed to be *the* structure of language. The notion of logical analysis advocated here does not amount to extracting something hidden; because there is nothing really hidden about language, we can see everything as it is. The only problem is to comprehend it. "Es ist eine Hauptquelle unseres Unverständnisses," Wittgenstein (1953, §122) says, "daß wir den Gebrauch unserer Wörter nicht *übersehen*. - Unserer Grammatik fehlt es an Übersichtlichkeit." And: "Die Philosophie stellt eben alles bloß hin, und erklärt und folgert nichts. - Da alles offen daliegt, ist auch nichts zu erklären." (ibid., §126)[182]. Thus, we propose to see logic and logical analysis not in terms of prospecting for *the* structure within the depths of language, but rather in terms of building watchtowers over language to help us comprehend language and see *a* structure[183].

---

[181]See Wittgenstein's (1953, §122) concept of *übersichtliche Darstellung*; and cf. Peregrin (1992).

[182]"A main source of our failure to understand is that we do not *command a clear view* of the use of words. Our grammar is lacking in this sort of perspicuity. ... Philosophy simply puts everything before us, and neither explains nor deduces anything. - Since everything lies open to view, there is nothing to explain."

[183]This is essentially the view envisaged by Quine (1972, p.451): "I find the phrase 'logical analysis' misleading, in its suggestion that we are exposing a logical structure that lay hidden in the sentence all along. ... When we move from verbal sentences to logical formulas we are merely retreating to a notation that has certain technical advantages, algorithmic and conceptual."

Thus, doing logical analysis we build a formal reconstruction of language which simplifies the real functioning of language in much the same way in which we build a simplified model of a complicated engine to demonstrate its principles of operation. As Stekeler-Weithofer (1986a, p.19) puts it:

> ...*jeder* Begriff der logischen Form und *jede* Semantik (...) bestenfalls angemessene *'Bilder'* (Vergleichsobjekte) sind, die wir zu bestimmten Zwecken, etwa der Erläuterung, Übersicht, der Einführung von Schreibweisen usw. - *neben* die schon im Gebrauch befindliche, uns vertraute und in ihren wesentlichen Teilen schon völlig *funktionsfähige* Sprache setzen.[184]

Entities posited in the course of such a reconstruction, such as models or possible worlds, may be of a great use; it is, however, not warranted to promote them to a more authentic reality than the entities the reconstruction of which brought them into being in the first place. To talk, e.g., about possibility and necessity in terms of possible worlds is quite legitimate and also illuminating, and it can help us to *understand* necessity and possibility; it is, however, not the only way to talk about these concepts and it is thus quite unwarranted to believe that something is possible *because* it is true in a possible world.

We try to understand truth, necessity or meaning, and therefore we take pains to carry out various reconstructions (such as, e.g. the model-theoretic reconstruction of the class of all necessary truths). The reconstructions can contribute to our understanding, but they are not explications in the sense that they would mediate the reduction of the analyzed concepts to simpler or more primitive ones, they are explications in the sense that they bring about criteria. The error of some theoreticians is that they are somewhat convinced that such concepts as truth and necessity *must* be reduced to something simpler. This is what was stressed in the conclusion of Wittgenstein's *Philosophical Investigations* (1953, §654):

---

[184]... any concept of logical form and any theoretical semantics develop at most pictures as objects of comparison. We set these models beside a linguistic usage we already are acquainted with. We know that in essential respects the language in use already fulfils its function properly. The juxtaposition of formal models serves special interests: to give an explication of problematic usages or an overview over possible usages or an introduction to special ways of expressions etc.

Unser Fehler ist dort nach einer Erklärung zu suchen, wo wir die Tatsachen als "Urphänomene" sehen sollten. D.h., wo wir sagen sollten: *dieses Sprachspiel wird gespielt.*[185]

Formal theories, and, in fact, theories in general, are *pictures*; they explain to the extent to which they mediate insight, no more. A formal theory is neither to replace, nor to rectify that which it is the theory of. The calculi of formal logic are not substitutes for natural language; nor do they provide an absolute norm governing the rational usage of natural language. They reside *beside* language and help us see its regularities, its "principles" - by way of comparison. This does not mean that they are not important - on the contrary, what they can do is to facilitate understanding, and understanding is the very thing our reason seeks.

[185]Our mistake is to look for an explanation where we ought to look at what happens as a "proto-phenomenon". That is, where we ought to have said: *this language-game is played.*

REFERENCES

Ambrose, A., ed. (1979): *Wittgenstein's Lectures, Cambridge, 1932-1935*, Blackwell, Oxford.

Austin, J.L. (1961): *Philosophical Papers* (ed. by J.O.Urmson and G.J.Warnock), Oxford University Press, London.

Bach, E. (1986): 'Natural Language Metaphysics', in *Logic, Methodology and Philosophy of Science 7* (ed. by R.B.Marcus, G.J.W.Dorn and P.Weingartner), North-Holland, Amsterdam, pp. 573-595.

Bach, E., Jelinek, E., Kratzer, A. and Partee, B., eds. (to appear): *Quantification in Natural Languages*, Kluwer, Dordrecht.

Barrett, R.B. and Gibson, R.F., eds. (1990): *Perspectives on Quine*, Blackwell, Oxford.

Barwise J. and Perry, J. (1983): *Situations and Attitudes*, MIT Press, Cambridge (Mass.).

Bäuerle, R. and Cresswell, M.J. (1989): 'Propositional Attitudes', in Gabbay and Guenthner.

Bealer, G. (1993): 'Universals', *Journal of Philosophy* 90, pp. 5-32.

Beth, E.W. (1955): 'Semantic Entailment and Formal Derivability', *Mededelingen van de Koninklijke Nederlandse Akademie van Wettenschappen, Afdeling Letterkunde*, N.R. Vol. 18, Amsterdam; reprinted in Hintikka (1969), pp. 9-41.

Bourbaki, N. (1958): *Élements de Mathématique I: Théorie des ensembles*, Hermann, Paris.

Brandom, R. (1983): 'Asserting', *Noûs* 17, pp. 637-650.

Brandom, R. (1985): 'Varieties of Understanding', in *Reason and Rationality in Natural Science* (ed. by N.Rescher), University Press of America, Lanham, pp. 27-51.

Brouwer, L.E.J. (1907): *Over de grodnslagen der wiskunde*, Thesis, Amsterdam; reprinted in and quoted from Brouwer: *Collected Works I: Philosophy and Foundations of Mathematics* (ed. A.Heyting), North-Holland, Amsterdam, 1975.

Carnap, R. (1931): 'Überwindung der Metaphysik durch logische Analyse der Sprache', *Erkenntnis* 2, pp. 219-241.

Carnap, R. (1934): 'On the Character of Philosophic Problems', *Philosophy of Science* 1, pp. 5-19.

Carnap, R. (1950a): 'Empiricism, Semantics and Ontology', *Revue International de Philosophie* 4, pp. 20-40.

Carnap, R. (1950b): *Logical Foundations of Probability*, University of Chicago Press, Chicago.

Carnap, R. (1955): 'Meaning and Synonymy in Natural Languages', *Philosophical Studies* 7, pp. 33-47.

Church, A. (1940): 'A Formulation of the Simple Theory of Types', *Journal of Symbolic Logic* 5, pp. 56-68.

Cresswell, M.J. (1982): 'The Autonomy of Semantics', in *Processes, Beliefs and Questions* (ed. by S.Peters and E.Saarinen), Reidel, Dordrecht, pp. 69-86.

Curry, H.B. (1961): 'Some Logical Aspects of Grammatical Structure', in *The Structure of Language and its Mathematical Aspects* (Proceedings of Symposia in Applied Mathematics 12, ed. by R.Jakobson), American Mathematical Society, Providence (R.I).

Davidson, D. (1967a): 'The Logical Form of Action Sentences', in *The Logic of Decision and Action* (ed. N.Rescher), The University of Pittsburgh, Pittsburgh; reprinted in Davidson (1980), pp. 105-122.

Davidson, D. (1967b): 'Truth and Meaning', *Synthèse* 17; reprinted in and quoted from Davidson (1984), pp. 17-36.

Davidson, D. (1970): 'Semantics for Natural Languages', *Linguaggi nella Società e nella Tecnica*, Edizioni di Comunità; reprinted in and quoted from Davidson (1984), pp. 55-64.

Davidson, D. (1973): 'In Defense of Convention T' in *Truth, Syntax and Modality* (ed. by H.Leblanc), North-Holland, Amsterdam; reprinted in and quoted from Davidson (1984), pp. 65-75.

Davidson, D. (1974): 'On the Very Idea of a Conceptual Scheme', *Proceedings and Addresses of the American Philosophical Association* 47; reprinted in and quoted from Davidson (1984), pp. 183-198.

Davidson, D. (1977a): 'Reality Without Reference', *Dialectica* 31; reprinted in and quoted from Davidson (1984), pp. 215-226.

Davidson, D. (1977b): 'The Method of Truth in Metaphysics', in *Midwest Studies in Philosophy 2: Studies in the Philosophy of Language* (ed. by P.A.French, T.E. Uehling, Jr., and H.K. Wettstein), University of Minnesota Press, Minneapolis; reprinted in and quoted from Davidson (1984), pp. 215-226.

Davidson, D. (1979): 'The Inscrutability of Reference', *Southwestern Journal of Philosophy* 10; reprinted in and quoted from Davidson (1984), pp. 227-241.

Davidson, D. (1980): *Essays on Actions and Events*, Clarendon Press, Oxford.

Davidson, D. (1984): *Inquiries into Truth and Interpretation*, Clarendon Press, Oxford.

Davidson, D. (1990): 'The Structure and Contents of Truth', *Journal of Philosophy* 87, pp. 279-328.

Davidson, D. and Harman, G., eds. (1972): *Semantics of Natural Language*, Reidel, Dordrecht.

Dennett, D.C. (1991): *Consciousness Explained*, Penguin, London.

Dummett, M. (1959): 'Truth', *Proceedings of the Aristotelian Society (New Series)* 59; reprinted in and quoted from Dummett (1978), pp. 1-24.

Dummett, M. (1976): 'What is a theory of meaning?' (II), in Evans and McDowell, pp. 67-137.

Dummett, M. (1978): *Truth and other Enigmas*, Duckworth, London.

Dummett, M. (1988): 'The Origins of Analytical Philosophy' I-II, *Lingua e Stile* 23, pp. 3-49, 171-210.

Etchemendy, J. (1988): 'Tarski on Truth and Logical Consequence', *Journal of Symbolic Logic* 53, pp. 51-79.

Etchemendy, J. (1990): *The Concept of Logical Consequence*, Harvard University Press, Cambridge (Mass.).

Evans, G. and McDowell, J., eds. (1976): *Truth and Meaning*, Oxford University Press, Oxford.

Feyerabend, P. (1978): *Science in a Free Society*, New Left Books, London.

Field, H. (1972): 'Tarski's Theory of Truth', *Journal of Philosophy* 69, pp. 347-375.

Fodor, J. (1987): *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*, MIT Press, Cambridge (Mass.).

Foucault, M. (1966): *Les Mots et les Choses*, Gallimard, Paris.

Frege, G. (1879): *Begriffsschrift*, Nebert, Halle; translated as *Begriffsschrift*, in Heijenoort (1971), pp. 1-82.

Frege, G. (1884): *Grundlagen der Arithmetik*, Koebner, Breslau; translated as *Foundations of Arithmetics*, Blackwell, Oxford, 1953.

Frege, G. (1885): *Über formale Theorien der Aritmetik*, Fischer, Jena.

Frege, G. (1892): 'Über Sinn und Bedeutung', *Zeitschrift für Philosophie und philosophische Kritik* 100, pp. 25-50.

Frege, G. (1892): 'Über Begriff und Gegenstand', *Vierteljahrschrift für wissentschaftliche Philosophie* 16, pp. 192-205.

Gabbay, D. and Guenthner, F., eds. (1989): *Handbook of Philosophical Logic IV*, Reidel, Dordrecht.

Gallin, D. (1975): *Intensional and Higher-order Modal Logic*, North-Holland, Amsterdam.

Gibson, R.F., jr. (1982): *The Philosophy of W.V.O.Quine*, University Presses of Florida, Tampa.

Gödel, K. (1930): 'Die Vollständigkeit der Axiome des logischen Funktionenkalküls', *Monatshefte für Mathematik und Physik* 37, pp. 349-360.

Goldfarb, W.D. (1979): 'Logic in the Twenties: the Nature of the Quantifier', *Journal of Symbolic Logic* 44, pp. 351-368.

Goodman, N. (1960): 'The Way the World is', *Review of Metaphysics* 14, pp. 48-56.

Goodman, N. (1978): *Ways of Worldmaking*, Harvester Press, Hassocks.

Goodman, N. (1984): *Of Mind and Other Matters*, Harvard University Press, Cambridge (Mass.).

Grice, P. (1989): *Studies in the Ways of Words*, Harvard University Press, Cambridge (Mass.).

Hacking, I. (1979): 'What is logic?', *Journal of Philosophy* 76, pp. 285-319.

Hacking, I. (1985): 'Styles of Scientific Reasoning', in Rajchmann and West, pp. 145-165.

Hagège, C. (1993): *The Language Builder: an Essay on the Human Signature in Linguistic Morphogenesis*, Benjamins, Amsterdam.

Hailperin, T. (1992): 'Herbrand Semantics, the Potential Infinite, and Ontology-Free Logic', *History and Philosophy of Logic* 13, pp. 69-90.

Hanson, N.R (1961): *Patterns of Discovery*, Cambridge University Press, Cambridge.

Heijenoort, J. van (1967): 'Logic as Calculus and Logic and Language', *Synthèse* 17, pp. 324-330.

Heijenoort, J. van, ed. (1971): *From Frege to Gödel: A Source Book from Mathematical Logic*, Harvard University Press, Cambridge (Mass.).

Hempel, C.G. (1935): 'On the Logical Positivists Theory of Truth', *Analysis* 2, pp. 49-59.

Henkin, L. (1949): 'The Completness of the First-order Functional Calculus', *Journal of Symbolic Logic* 14, pp. 159-166.

Henkin, L. (1950): 'Completness in the Theory of Types', *Journal of Symbolic Logic* 15, pp. 81-91.

Hilbert, D. (1925): 'Über das Unendliche', *Mathematische Annalen* 95, pp. 161-190; translated as 'On the Infinite', in Heijenoort (1971), pp. 367-392.

Hintikka, J., ed. (1969): *The Philosophy of Mathematics*, Oxford University Press, Oxford.

Hintikka, J. (1978): 'Impossible Possible Worlds Vindicated', in *Game-Theoretical Semantics* (ed. by E.Saarinen), Reidel, Dordrecht, pp. 367-379.

Hintikka, J. (1984): 'A Hundred Years Later: the Rise and Fall of Frege's Influence in Language Theory', *Synthèse* 59, pp. 27-49.

Hintikka, J. (1990): 'Quine as a Member of the Tradition of the Universality of Language', in Barrett and Gibson, pp. 159-175.

Hintikka, M.B. and Hintikka, J. (1986): *Investigating Wittgenstein*, Blackwell, Oxford.

Horwich, P. (1990): *Truth*, Blackwell, Oxford.

Horwich, P. (ms.): 'What is it like to be a deflationary theory of meaning?', manuscript.

Husserl, E. (1900/1901): *Logische Untersuchungen*, Niemeyer, Halle.

Husserl, E. (1977): *Cartesianische Meditationen*, Meiner, Hamburg.

Janssen, T.M.V. (1983): *Foundations and Applications of Montague Grammar*, dissertation, Mathematical Centre, Amsterdam.

Kamp, H. (1981): 'A Theory of Truth and Semantic Representation', in *Formal Methods in the Study of Language* (ed. by J.Groenendijk, T.Janssen and M.Stokhof), Mathematical Centre, Amsterdam; reprinted in *Truth, Interpretation and Information* (ed. by J.Groenendijk, M.Stokhof and T.Janssen), Foris, Dordrecht, 1984, pp. 1-41.

Kemeny, J.G. (1948): 'Models of Logical Systems', *Journal of Symbolic Logic* 13, pp. 16-30.

Kenny, A. (1972): *Wittgenstein*, Penguin, London.

Kracht, M. (1993): 'How Completeness and Correspondence Theory Got Married', in *Diamonds and Defaults* (ed. by M.de Rijke), Kluwer, Dordrecht, pp. 175-214.

Kreisel, G. (1967): 'Informal Rigour and Completeness Proofs', in *Problems in the Philosophy of Mathematics* (ed. by I.Lakatos), North-Holland, Amsterdam; reprinted in and quoted from Hintikka (1969), pp. 78-94.

Kripke, S. (1963): 'Semantical Considerations on Modal Logic', *Acta Philosophica Fennica* 16, pp. 83-94.

Kripke, S. (1972): 'Naming and Necessity', in Davidson and Harman, pp. 253-355.

Kripke, S. (1976): 'Is There a Problem about Substitutional Quantification?', in Evans and McDowell, pp. 325-419.

Leilich, J. (1983): *Die Autonomie Der Sprache. Ein Grundgedanke Wittgensteins*, Profil, München.

Lemmon, E. (1965): *Beginning Logic*, Nelson, London.

Lemmon, E. (1966): 'Algebraic Semantics for Modal Logic', *Journal of Symbolic Logic* 31, pp. 46-65 and 191-218.

Locke, J. (1690): *An Essay Concerning Human Understanding*, London.

Lorenzen, P. (1957): 'Das Aktual-Unendliche in der Mathematik', *Philosophia Naturalis* 4; reprinted in Lorenzen: *Methodisches Denken*, Suhrkamp, Frankfurt, 1988, pp. 94-103.

Lukasiewicz, J. and Tarski, A. (1930): 'Untersuchungen Über den Aussagenkalkül', *Comptes Rendus des Seances de la Societe des Sciences et des Lettres de Varsovie* 23, pp. 30-50.

Malcolm, N. (1940): 'Are Necessary Propositions Really Verbal?', *Mind* 49, pp. 189-203.

Mates, B. (1968): 'Leibniz on Possible Worlds', in *Logic, Methodology and Philosophy of Science* 3 (ed. by B. Van Rootsellar and J.F.Stall), North Holland, Amsterdam.

McGee, V. (1992): 'Two Problems with Tarski's Theory of Consequence', *Proceedings of the Aristotelian Society (New Series)* 92, pp. 273-292.

McGuinness, B. and Olivieri, G., eds. (1994): *The Philosophy of Michael Dummett*, Kluwer, Dordrecht.

Montague, R. (1974): *Formal Philosophy: selected papers of R.Montague* (ed. by R.Thomason), Yale University Press, New Haven.

Morris, C.W. (1966): *Foundations of the Theory of Signs* (Internatilnal Encyclopedia of Unified Science 1), University of Chicago Press, Chicago.

Mostowski, A. (1965): 'Thirty Years of Foundational Studies', *Acta Philosophica Fennica* 17.

Parsons, T. (1990): *Events in the Semantics of English*, MIT Press, Cambridge (Mass.).

Pears, D. (1994): 'Philosophical Theorizing and Particularism: Michael Dummett on Wittgenstein's Later Philosophy of Language', in McGuinness and Olivieri, pp. 45-57.

Penco, C. (1994): 'Dummett and Wittgenstein's Philosophy of Mathematics', in McGuinness and G.Olivieri, pp. 113-136.

Peregrin, J. (1992): 'Sprache und ihre Formalisierung', *Deutsche Zeitschrift für Philosophie* 40, pp. 237-244.

Peregrin, J. (1993a): 'Possible Worlds: a Critical Analysis', *Prague Bulletin of Mathematical Linguistics* No. 59-60, pp. 9-21.

Peregrin, J. (1993b): 'Language and Formalisms', *Linguistica Pragensia* No. 2, pp. 74-80.

Peregrin, J. (1993c): 'Is Language a Code?', *From the Logical Point of View* No. 2, pp. 73-79.

Peregrin, J. (1994): 'Interpreting Formal Logic', *Erkenntnis* 40, pp. 5-20.

Peregrin, J. (1995): 'Structuralism: Slogan or Concept?', *Lingua e Stile* 30, 475-495.

Peregrin, J. (forthcoming a): 'Structural Linguistics and Formal Semantics', to apear in *Travaux de Circle Linguistique de Prague*.

Peregrin, J. (forthcoming b): 'Structure and Meaning', to apear in *Semiotica*.

Peregrin, J. (fortcoming c): 'Philosophy and the Pursuit of Meaning', sumbitted to publication.

Peregrin, J. (forthcoming d): 'Language and its Models', submitted to publication.

Putnam, H. (1975): 'The Meaning of "Meaning"', in *Language, Mind and Knowledge* (Studies in the Philosophy of Science VII, ed. by K.Gunderson), University of Minnesota Press, Minneapolis; reprinted in Putnam: *Mind, Language and Reality (Philosophical Papers 2)*, Cambridge University Press, Cambridge, pp. 215-271.

Putnam, H. (1981): *Reason, Truth and History*, Cambridge University Press, Cambridge.

Putnam, H. (1984): 'After Ayer, After Empiricism', *Partisan Review* 51; reprinted as 'After Empiricism' in Rajchmann and West, pp. 21-30.

Putnam, H. (1985): 'A Comparison of Something with Something Else', *New Literary History* 17, pp. 61-79.

Putnam, H. (1988): *Representation and Reality*, MIT Press, Cambridge (Mass.)

Putnam, H. (1990): *Realism with a Human Face*, Harvard University Press, Cambridge (Mass.).

Quine, W.V.O. (1960a): *Word and Object*, MIT Press, Cambridge (Mass.).

Quine, W.V.O. (1960b): 'Variables explained away', *Proceedings of the American Philosophical Society* 104, pp. 343-347.

Quine, W.V.O. (1963): *From the Logical Point of View*, Harper and Row, New York.

Quine, W.V.O. (1969): *Ontological Relativity and Other Essays*, Columbia University Press, New York.

Quine, W.V.O. (1963): *Set Theory and its Logic* (revised edition), Belnap Press, Cambridge (Mass.).

Quine, W.V.O. (1972): 'Methodological Reflections on Current Linguistic Theory', in Davidson and Harman, pp. 442-454.

Quine, W.V.O. (1980): 'Grammar, Truth and Logic', in *Philosophy and Grammar* (ed. by S.Kanger and S.Öhman), Reidel, Dordrecht, pp. 17-28.

Quine, W.V.O. (1986): 'Reply to D.Føllesdal', in *Philosophy of W.V.O.Quine* (ed. by L.E.Hahn and P.A.Schilpp), Open Court, La Salle, pp. 114-115.

Quine, W.V. (1991): 'Two Dogmas in Retrospect', *Canadian Journal of Philosophy* 21, pp. 265-274.

Quine, W.V.O. (1992): 'Structure and Nature', *Journal of Philosophy* 89, pp. 5-9.

Quine, W.V.O. (1994): 'Response to Bergström', *Inquiry* 37, pp. 496-497.

Rajchmann, J. and West, C., eds. (1985): *Post-Analytic Philosophy*, Columbia University Press, New York

Robinson, A. (1965): 'Formalism 64', in *Proceedings of the LMPS Congress*, Amsterdam; reprinted in Robinson: *Selected Papers*, New Haven, 1979, pp. 505-523.

Romanos, G.D. (1983): *Quine and Analytic Philosophy*, MIT Press, Cambridge (Mass.).

Rorty, R. (1989): *Contingency, Irony and Solidarity*, Cambridge University Press, Cambridge.

Ros, A. (1989): *Begründung und Begriff* I, Meiner, Hamburg.

Ross, W.D., ed. (1930): *The Works of Aristotle*, Vol. 2, Oxford.

Russell, B. (1905): 'On denoting', *Mind* 14, pp. 479-493.

Russell, B. (1914): *Our Knowledge of the External World*, Allen and Unwin, London.

Sapir, E. (1921): *Language; an Introduction to the Study of Speech*, Harcourt, New York.

Saussure, F. de (1931): *Cours de linguistique générale*, Payot, Paris; translated as *Course in General Linguistics*, Philosophical Library, New York, 1959.

Schiffer, S. (1972): *Meaning*, Clarendon Press, Oxford.

Schlick, M. (1932): 'The Future of Philosophy', *College of the Pacific Publications in Philosophy* 1, pp. 45-62.

Searle, J. (1983): *Intentionality*, Cambridge University Press, Cambridge.

Sellars, W. (1956): 'The Myth of the Given: Three Lectures on Empiricism and the Philosophy of Mind', in *The Foundations of Science and the Concepts of Psychology and Psychoanalysis* (Minnesota Studies in the Philosophy of Science I, ed. by H.Feigl and M.Scriven), University of Minnesota Press, Minneapolis; reprinted in and quoted from Sellars (1963b).

Sellars, W. (1963a): 'Empiricism and Abstract Entities', in *The Philosophy of Rudolf Carnap* (ed. P.A.Schilpp), Open Court, LaSalle, pp. 431-468.

Sellars, W. (1963b): *Science, Perception and Reality*, Routledge, New York.

Sgall, P., Hajičová, E. and Panevová, J. (1986): *The Meaning of the Sentence in its Semantic and Pragmatic Aspects*, Academia, Prague.

Sgall, P. and Peregrin, J. (forthcoming): 'Meaning and "Propositional Attitudes"', to appear in *In the World of Signs* (Festschrift for J.Pelc).

Shoenfield, J.R. (1967): *Mathematical Logic*, Addison-Wesley, Reading.

Skolem, T.: 'Begründung der elementaren Arithmetik durch die rekurrierende Denkweise ohne Anwendung scheinbarer Veränderlichen mit unendlichem Ausdehnungsbereich', *Videnkapsselskapets skrifter. I. Matematisk-naturvidenskabelig klasse* No. 6; translated as 'The foundations of elementary arithmetic established by means of the recursive mode of thought, without the use of apparent variables ranging over infinite domains' in Heijenoort (1967), pp. 302-333.

Skyrms, B. (1993): 'Logical Atoms and Combinatorial Possibility', *Journal of Philosophy* 90, pp. 219-232.

Soames, S. (1984): 'What is a Theory of Truth?', *Journal of Philosophy* 81, pp. 411-429.

Steck, M., ed. (1941): *Unbekannte Briefe Frege's über die Grundlagen der Geometrie und Antwortbrief Hilbert's an Frege*, Komissionsverlag der Weiß'schen Universitätsbuchhandlung Heidelberg, Heidelberg.

Stekeler-Weithofer, P. (1986a): *Grundprobleme der Logik*, de Gruyter, Berlin.

Stekeler-Weithofer, P. (1986b): *Sinnkriterien* I, Universität Konstanz.

Stekeler-Weithofer, P. (1992): 'Plato and the Method of Science', *History of Philosophy Quarterly* 9, pp. 359-378.

Tarski, A. (1932): 'Der Wahrheitsbegriff in den Sprachen der deduktiven Disziplinen', *Akademischer Anzeiger der Akademie der Wissenschaften in Wien* 69; translated as 'The Concept of Truth in Formalized Languages' in Tarski (1956), pp. 152-278.

Tarski, A. (1936): 'Über den Begriff der logischen Folgerung', *Actes du Congrés International de Philosophique Scientifique* 7, pp. 1-11; translated as 'On the Concept of Logical Consequence', in Tarski (1956), pp. 409-420.

Tarski, A. (1944): 'The Semantic Conception of Truth', *Philosophy and Phenomenological Research* 4, pp. 341-375.

Tarski, A. (1956): *Logic, Semantics, Metamathematics*, Clarendon Press, Oxford.

Tichý, P. (1988): *The Foundations of Frege's Logic*, de Gruyter, Berlin.

Tichý, P. (1992): 'The Scandal of Linguistics', *From the Logical Point of view* No. 3, pp. 70-80.

Vopěnka, P. (1979): *Mathematics in the Alternative Set Theory*, Teubner, Leipzig.

Whorf, B.L. (1956): *Language, Thought and Reality*, MIT Press, Cambridge (Mass.).

Wittgenstein, L. (1922): *Tractatus Logico-Philosophicus*, Routledge, London; translated as *Tractatus Logico-Philosophicus*, Routledge, London, 1961.

Wittgenstein, L. (1953): *Philosophische Untersuchungen*, Blackwell, Oxford; translated as *Philosophical Investigations*, Blackwell, Oxford, 1953.

Wittgenstein, L. (1956): *Bemerkungen über die Grundlagen der Mathematik*, Blackwell, Oxford; translated as *Remarks on the Foundations of Mathematics*, Blackwell, Oxford, 1956.

Wittgenstein, L. (1964): *Philosophische Bemerkungen*, Suhrkamp, Frankfurt; translated as *Philosophical Remarks*, Blackwell, Oxford, 1975.

Wittgenstein, L. (1969): *Philosohpische Grammatik*, Suhrkamp, Frankfurt; translated as *Philosophical Grammar*, Blackwell, Oxford, 1974.

Wittgenstein, L. (1977): *Vermischte Bemerkungen* (ed. by G.H. von Wright), Suhrkamp, Frankfurt; translated as *Culture and Value*, Blackwell, Oxford, 1980.

NAME INDEX

SUBJECT INDEX

174. P. Ziff, *Antiaesthetics.* An Appreciation of the Cow with the Subtile Nose. 1984 ISBN 90-277-1773-7
175. W. Balzer, D. A. Pearce, and H.-J. Schmidt (eds.), *Reduction in Science.* Structure, Examples, Philosophical Problems. 1984 ISBN 90-277-1811-3
176. A. Peczenik, L. Lindahl and B. van Roermund (eds.), *Theory of Legal Science.* Proceedings of the Conference on Legal Theory and Philosophy of Science (Lund, Sweden, December 1983). 1984 ISBN 90-277-1834-2
177. I. Niiniluoto, *Is Science Progressive?* 1984 ISBN 90-277-1835-0
178. B. K. Matilal and J. L. Shaw (eds.), *Analytical Philosophy in Comparative Perspective.* Exploratory Essays in Current Theories and Classical Indian Theories of Meaning and Reference. 1985 ISBN 90-277-1870-9
179. P. Kroes, *Time: Its Structure and Role in Physical Theories.* 1985 ISBN 90-277-1894-6
180. J. H. Fetzer, *Sociobiology and Epistemology.* 1985 ISBN 90-277-2005-3; Pb 90-277-2006-1
181. L. Haaparanta and J. Hintikka (eds.), *Frege Synthesized.* Essays on the Philosophical and Foundational Work of Gottlob Frege. 1986 ISBN 90-277-2126-2
182. M. Detlefsen, *Hilbert's Program.* An Essay on Mathematical Instrumentalism. 1986 ISBN 90-277-2151-3
183. J. L. Golden and J. J. Pilotta (eds.), *Practical Reasoning in Human Affairs.* Studies in Honor of Chaim Perelman. 1986 ISBN 90-277-2255-2
184. H. Zandvoort, *Models of Scientific Development and the Case of Nuclear Magnetic Resonance.* 1986 ISBN 90-277-2351-6
185. I. Niiniluoto, *Truthlikeness.* 1987 ISBN 90-277-2354-0
186. W. Balzer, C. U. Moulines and J. D. Sneed, *An Architectonic for Science.* The Structuralist Program. 1987 ISBN 90-277-2403-2
187. D. Pearce, *Roads to Commensurability.* 1987 ISBN 90-277-2414-8
188. L. M. Vaina (ed.), *Matters of Intelligence.* Conceptual Structures in Cognitive Neuroscience. 1987 ISBN 90-277-2460-1
189. H. Siegel, *Relativism Refuted.* A Critique of Contemporary Epistemological Relativism. 1987 ISBN 90-277-2469-5
190. W. Callebaut and R. Pinxten, *Evolutionary Epistemology.* A Multiparadigm Program, with a Complete Evolutionary Epistemology Bibliograph. 1987 ISBN 90-277-2582-9
191. J. Kmita, *Problems in Historical Epistemology.* 1988 ISBN 90-277-2199-8
192. J. H. Fetzer (ed.), *Probability and Causality.* Essays in Honor of Wesley C. Salmon, with an Annotated Bibliography. 1988 ISBN 90-277-2607-8; Pb 1-5560-8052-2
193. A. Donovan, L. Laudan and R. Laudan (eds.), *Scrutinizing Science.* Empirical Studies of Scientific Change. 1988 ISBN 90-277-2608-6
194. H.R. Otto and J.A. Tuedio (eds.), *Perspectives on Mind.* 1988 ISBN 90-277-2640-X
195. D. Batens and J.P. van Bendegem (eds.), *Theory and Experiment.* Recent Insights and New Perspectives on Their Relation. 1988 ISBN 90-277-2645-0
196. J. Österberg, *Self and Others.* A Study of Ethical Egoism. 1988 ISBN 90-277-2648-5
197. D.H. Helman (ed.), *Analogical Reasoning.* Perspectives of Artificial Intelligence, Cognitive Science, and Philosophy. 1988 ISBN 90-277-2711-2

198. J. Wolenski, *Logic and Philosophy in the Lvov-Warsaw School.* 1989 ISBN 90-277-2749-X
199. R. Wójcicki, *Theory of Logical Calculi.* Basic Theory of Consequence Operations. 1988 ISBN 90-277-2785-6
200. J. Hintikka and M.B. Hintikka, *The Logic of Epistemology and the Epistemology of Logic.* Selected Essays. 1989 ISBN 0-7923-0040-8; Pb 0-7923-0041-6
201. E. Agazzi (ed.), *Probability in the Sciences.* 1988 ISBN 90-277-2808-9
202. M. Meyer (ed.), *From Metaphysics to Rhetoric.* 1989 ISBN 90-277-2814-3
203. R.L. Tieszen, *Mathematical Intuition.* Phenomenology and Mathematical Knowledge. 1989 ISBN 0-7923-0131-5
204. A. Melnick, *Space, Time, and Thought in Kant.* 1989 ISBN 0-7923-0135-8
205. D.W. Smith, *The Circle of Acquaintance.* Perception, Consciousness, and Empathy. 1989 ISBN 0-7923-0252-4
206. M.H. Salmon (ed.), *The Philosophy of Logical Mechanism.* Essays in Honor of Arthur W. Burks. With his Responses, and with a Bibliography of Burk's Work. 1990 ISBN 0-7923-0325-3
207. M. Kusch, *Language as Calculus vs. Language as Universal Medium.* A Study in Husserl, Heidegger, and Gadamer. 1989 ISBN 0-7923-0333-4
208. T.C. Meyering, *Historical Roots of Cognitive Science.* The Rise of a Cognitive Theory of Perception from Antiquity to the Nineteenth Century. 1989 ISBN 0-7923-0349-0
209. P. Kosso, *Observability and Observation in Physical Science.* 1989 ISBN 0-7923-0389-X
210. J. Kmita, *Essays on the Theory of Scientific Cognition.* 1990 ISBN 0-7923-0441-1
211. W. Sieg (ed.), *Acting and Reflecting.* The Interdisciplinary Turn in Philosophy. 1990 ISBN 0-7923-0512-4
212. J. Karpiński, *Causality in Sociological Research.* 1990 ISBN 0-7923-0546-9
213. H.A. Lewis (ed.), *Peter Geach: Philosophical Encounters.* 1991 ISBN 0-7923-0823-9
214. M. Ter Hark, *Beyond the Inner and the Outer.* Wittgenstein's Philosophy of Psychology. 1990 ISBN 0-7923-0850-6
215. M. Gosselin, *Nominalism and Contemporary Nominalism.* Ontological and Epistemological Implications of the Work of W.V.O. Quine and of N. Goodman. 1990 ISBN 0-7923-0904-9
216. J.H. Fetzer, D. Shatz and G. Schlesinger (eds.), *Definitions and Definability.* Philosophical Perspectives. 1991 ISBN 0-7923-1046-2
217. E. Agazzi and A. Cordero (eds.), *Philosophy and the Origin and Evolution of the Universe.* 1991 ISBN 0-7923-1322-4
218. M. Kusch, *Foucault's Strata and Fields.* An Investigation into Archaeological and Genealogical Science Studies. 1991 ISBN 0-7923-1462-X
219. C.J. Posy, *Kant's Philosophy of Mathematics.* Modern Essays. 1992 ISBN 0-7923-1495-6
220. G. Van de Vijver, *New Perspectives on Cybernetics.* Self-Organization, Autonomy and Connectionism. 1992 ISBN 0-7923-1519-7
221. J.C. Nyíri, *Tradition and Individuality.* Essays. 1992 ISBN 0-7923-1566-9
222. R. Howell, *Kant's Transcendental Deduction.* An Analysis of Main Themes in His Critical Philosophy. 1992 ISBN 0-7923-1571-5

223. A. García de la Sienra, *The Logical Foundations of the Marxian Theory of Value.* 1992 ISBN 0-7923-1778-5
224. D.S. Shwayder, *Statement and Referent.* An Inquiry into the Foundations of Our Conceptual Order. 1992 ISBN 0-7923-1803-X
225. M. Rosen, *Problems of the Hegelian Dialectic.* Dialectic Reconstructed as a Logic of Human Reality. 1993 ISBN 0-7923-2047-6
226. P. Suppes, *Models and Methods in the Philosophy of Science: Selected Essays.* 1993 ISBN 0-7923-2211-8
227. R. M. Dancy (ed.), *Kant and Critique: New Essays in Honor of W. H. Werkmeister.* 1993 ISBN 0-7923-2244-4
228. J. Woleński (ed.), *Philosophical Logic in Poland.* 1993 ISBN 0-7923-2293-2
229. M. De Rijke (ed.), *Diamonds and Defaults.* Studies in Pure and Applied Intensional Logic. 1993 ISBN 0-7923-2342-4
230. B.K. Matilal and A. Chakrabarti (eds.), *Knowing from Words.* Western and Indian Philosophical Analysis of Understanding and Testimony. 1994 ISBN 0-7923-2345-9
231. S.A. Kleiner, *The Logic of Discovery.* A Theory of the Rationality of Scientific Research. 1993 ISBN 0-7923-2371-8
232. R. Festa, *Optimum Inductive Methods.* A Study in Inductive Probability, Bayesian Statistics, and Verisimilitude. 1993 ISBN 0-7923-2460-9
233. P. Humphreys (ed.), *Patrick Suppes: Scientific Philosopher.* Vol. 1: Probability and Probabilistic Causality. 1994 ISBN 0-7923-2552-4
234. P. Humphreys (ed.), *Patrick Suppes: Scientific Philosopher.* Vol. 2: Philosophy of Physics, Theory Structure, and Measurement Theory. 1994 ISBN 0-7923-2553-2
235. P. Humphreys (ed.), *Patrick Suppes: Scientific Philosopher.* Vol. 3: Language, Logic, and Psychology. 1994 ISBN 0-7923-2862-0
     Set ISBN (Vols 233–235) 0-7923-2554-0
236. D. Prawitz and D. Westerståhl (eds.), *Logic and Philosophy of Science in Uppsala.* Papers from the 9th International Congress of Logic, Methodology, and Philosophy of Science. 1994 ISBN 0-7923-2702-0
237. L. Haaparanta (ed.), *Mind, Meaning and Mathematics.* Essays on the Philosophical Views of Husserl and Frege. 1994 ISBN 0-7923-2703-9
238. J. Hintikka (ed.), *Aspects of Metaphor.* 1994 ISBN 0-7923-2786-1
239. B. McGuinness and G. Oliveri (eds.), *The Philosophy of Michael Dummett.* With Replies from Michael Dummett. 1994 ISBN 0-7923-2804-3
240. D. Jamieson (ed.), *Language, Mind, and Art.* Essays in Appreciation and Analysis, In Honor of Paul Ziff. 1994 ISBN 0-7923-2810-8
241. G. Preyer, F. Siebelt and A. Ulfig (eds.), *Language, Mind and Epistemology.* On Donald Davidson's Philosophy. 1994 ISBN 0-7923-2811-6
242. P. Ehrlich (ed.), *Real Numbers, Generalizations of the Reals, and Theories of Continua.* 1994 ISBN 0-7923-2689-X
243. G. Debrock and M. Hulswit (eds.), *Living Doubt.* Essays concerning the epistemology of Charles Sanders Peirce. 1994 ISBN 0-7923-2898-1
244. J. Srzednicki, *To Know or Not to Know.* Beyond Realism and Anti-Realism. 1994 ISBN 0-7923-2909-0
245. R. Egidi (ed.), *Wittgenstein: Mind and Language.* 1995 ISBN 0-7923-3171-0
246. A. Hyslop, *Other Minds.* 1995 ISBN 0-7923-3245-8
247. L. Pólos and M. Masuch (eds.), *Applied Logic: How, What and Why.* Logical Approaches to Natural Language. 1995 ISBN 0-7923-3432-9

248. M. Krynicki, M. Mostowski and L.M. Szczerba (eds.), *Quantifiers: Logics, Models and Computation.* Volume One: Surveys. 1995 ISBN 0-7923-3448-5
249. M. Krynicki, M. Mostowski and L.M. Szczerba (eds.), *Quantifiers: Logics, Models and Computation.* Volume Two: Contributions. 1995 ISBN 0-7923-3449-3
Set ISBN (Vols 248 + 249) 0-7923-3450-7
250. R.A. Watson, *Representational Ideas from Plato to Patricia Churchland.* 1995 ISBN 0-7923-3453-1
251. J. Hintikka (ed.), *From Dedekind to Gödel.* Essays on the Development of the Foundations of Mathematics. 1995 ISBN 0-7923-3484-1
252. A. Wiśniewski, *The Posing of Questions.* Logical Foundations of Erotetic Inferences. 1995 ISBN 0-7923-3637-2
253. J. Peregrin, *Doing Worlds with Words.* Formal Semantics without Formal Metaphysics. 1995 ISBN 0-7923-3742-5

KLUWER ACADEMIC PUBLISHERS – DORDRECHT / BOSTON / LONDON