CLARENDON LIBRARY OF LOGIC AND PHILOSOPHY

George Bealer

Quality and Concept

Clarendon Press

This book has been printed digitally and produced in a standard specification in order to ensure its continuing availability

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship, and education by publishing worldwide in

Oxford New York

Auckland Bangkok Buenos Aires Cape Town Chennai
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata
Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi
São Paulo Shanghai Singapore Taipei Tokyo Toronto

Oxford is a registered trade mark of Oxford University Press in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

Reprinted 2002

ISBN 0-19-824428-2

Contents

. . . *logos* is born for us through the weaving together of forms.

Plato, *Sophist*

Introduction

1. The Call for a Theory

Properties, relations, and propositions permeate the world. They fix its logical, causal, and phenomenal order, and they structure our thoughts and words about it. So a theory of properties, relations, and propositions (PRPs) should unify a great many topics central to logic, metaphysics, psychology, and theory of language. A theory of PRPs also should handle all topics, including those in mathematics, previously thought to be the province of set theory. It is therefore ironic that by comparison with set theory the theory of PRPs has been relatively neglected over the past century. To be sure, significant advances have been made by Frege and Church, by Russell, and by Carnap and his followers. Yet in spite of the elegance and historical significance of these efforts, each suffers from many interrelated difficulties. Because these are often profound, it is little wonder that research on the theory of PRPs has been held back. The theory has never been adequately formulated.

In this work I try to develop a more nearly adequate formulation of the theory. I attempt to take a fresh look at the subject and to avoid preconceptions that shaped earlier efforts. I hope that my formulation can be applied easily and naturally in each of the above disciplines. To the extent that it can, this formulation will contrast sharply with some of its precursors, which are known for their excessive complication. Ease and naturalness in practice are, however, not the only tests that the theory must pass. At least as important is that it should come up to the high standard of rigor established by axiomatic set theory.

Two underlying tenets shape the work. The first is that properties, relations, and propositions are real, irreducible entities. In spite of the now dominant trend in modern mathematical thought, we should not be enticed into treating them as some special kind of set or function. If anything, sets and functions should be treated as surrogates for certain properties and relations. The second tenet is that the theory of PRPs, unlike set theory, is a full-fledged part of logic as traditionally conceived. So in the theory of PRPs we find a purely logical theory that is simultaneously a foundation for philosophy, psychology, theory of language, and mathematics. What emerges in the course of this work is a philosophy of logical realism, the view that solutions to foundational problems in metaphysics and science are to be found not in empiricism, naturalism, or idealism but rather in logic and, specifically, in a logic that embraces metaphysical realism.

2. Two Traditional Conceptions of Properties, Relations, and Propositions

Historically, there have been two fundamentally different conceptions of properties, relations, and propositions. On the first conception intensional entities are considered to be identical if and only if they are *necessarily equivalent*. So on this conception there are no constraints on what is to count as a correct definition beyond the requirement of necessary equivalence. For example, both of the following sentences taken from contemporary philosophy:[1]

(a) x is grue *iff* x is green if examined before t and blue otherwise.

(b) x is green *iff* x is grue if examined before t and bleen otherwise.

qualify as correct definitions on this conception.

On the second conception, by contrast, each definable intensional entity is such that, when it is defined completely, it has a *unique, non-circular definition*. (The possibility that such complete definitions might in some or even all cases be infinite need not be ruled out.) Hence, on this conception there are severe constraints on what is to count as a correct definition. For example, in view of its stipulative character, the original definition (a) of grue in terms of green (and blue) is certainly correct even if green should itself be definable. On the assumption that there is a unique way to spell out completely the correct definition of grue, it follows that, since green appears in *a* correct definition of grue, green must show up in *the* correct definition of grue as a defined or undefined term. Consequently, on the assumption that correct definitions cannot be circular, green cannot in turn be defined in terms of grue. Thus, although (a) and (b) above both express necessary truths, (a) alone is a correct definition on the second conception of intensional entities. Although necessary equivalence is a necessary condition for identity, it is not a sufficient condition.

Consider another example:

(c) x is a trilateral *iff* x is a closed plane figure having three sides.
(d) x is a trilateral *iff* x is a closed plane figure having three angles.

On the first conception both (c) and (d) count as correct definitions since they both express necessary truths. On the second conception, by contrast, (d) does not count as a correct definition; only (c) does. To see why, notice that because of the Latin roots of 'trilateral', (c) is surely correct. (This is not to say that it cannot be carried further.) Yet (c) and (d) are exactly alike except that 'angle' occurs in (d) where 'side' occurs in (c). Therefore, on the assumption that there is a unique way to spell out completely the definition of trilateral, (d) cannot also be a correct definition. for that would require the property of being an angle and the property of being a side to be identical. However, they are not even contingently equivalent; in fact, their instances are necessarily different.

The first conception of intensional entities is built into the possible-worlds treatment of PRPs. Indeed, this conception is commonly attributed to Leibniz; whether Leibniz actually subscribed to it, however, is open to doubt. This conception also underlies Alonzo Church's "Alternative (2)" formulation of Frege's theory of senses.[2]

The second conception of intensional entities has a far livelier history. Perhaps the clearest instance of it is to be found in Russell's doctrine of logical atomism. (On this doctrine it is required that all complete definitions be finite as well as unique and non-circular.) Traces of this conception are also evident in Leibniz's remarks on the distinction between simple and complex properties. Moreover, if concepts (ideas, thoughts) are identified with PRPs, evidence of this conception can be found in the writings of modern philosophers from Descartes and Locke, through Kant, and on to even Frege. Yet in spite of its lively history, this conception has to my knowledge never been invoked as the intuitive motivation for a formal theory of PRPs. Even though Russell's informal doctrine of logical atomism provides us with perhaps the clearest instance of this conception, *Principia Mathematica* itself is, ironically, neutral with regard to the two conceptions. And despite what one might expect, Alonzo Church does not intuitively motivate his "Alternative (0)" formulation of Frege's theory of senses with this conception of PRPs; instead, the intuitive motivation that Church explicitly invokes is a problematic conception of synonymy based on the notion of synonymous isomorphism.[3] A careful study of Church's axioms[4] reveals, however, that it is the second conception of PRPs that implicitly underlies this formulation of Frege's theory.

The first conception of PRPs is ideally suited for treating the modalities—necessity, possibility, impossibility, contingency, etc. However, it has proved to be of little value in the treatment of intentional matters—belief, desire, perception, decision, assertion, etc. Indeed, it has led its major contemporary proponents to construct theories that provide strikingly inadequate treatments of them. The second conception, on the other hand, while ideally suited for the treatment of intentional matters,[5] has only complicated the treatment of the modalities. The relevance of the first conception to modality and the relevance of the second conception to intentionality suggest that we should, at least provisionally, develop both conceptions side by side. This dual approach has a special advantage. For between these two traditional conceptions there are any number of intermediate conceptions, and one should leave open the possibility that in actual contexts of thought and speech any of these conceptions might be at work. What is unique about the two traditional conceptions is that they determine limits between which all other natural conceptions conveniently fall. If the two traditional conceptions can be successfully formulated, then it will be a straightforward affair to adapt the resulting theories to capture any of the intermediate conceptions.[6] So a dual approach that succeeds in capturing these two conceptions may be thought of as a prism that indirectly captures the entire spectrum of intermediate conceptions. The success of the dual approach, thus, does not ride on the correctness of either traditional conception; rather it rides only on the correctness of some conception or other lying in this spectrum.

The value of both traditional conceptions of PRPs is evident. Therefore, in what follows I propose to develop both conceptions side by side. Along the way I will treat each topic in our general subject area—intensionality, predication, class, number, meaning, truth, necessity, analyticity, intentionality, and consciousness. Near the end of this work this dual development will culminate in a synthesis of the two conceptions. The result will be a unified theory of qualities and concepts. It is in terms of this theory that non-circular definitions of truth, necessity, analyticity, intentionality, and consciousness will be framed. However, before I get too far ahead of myself, allow me to sketch in some of the details.

3. Preview of the Theory of Qualities and Concepts

The work is divided into three parts. Briefly, the first part builds a complete logic for modal and intentional matters. This intensional logic provides a foundation for the subsequent study of PRPs. Then in the second part this intensional logic is extended by the addition of the fundamental logical relation of predication. This forces the theory of PRPs to look for the first time upon the spectre of incompleteness and paradox. Yet the theory also benefits from this extension, for now it is able to provide a natural account of number independent of artificial set-theoretical constructions. Finally, unification is sought in the third part. To start with, intensional logic is derived within an extensional logic. Next the semantic theory of truth and the theory of meaning are both constructed within a semantical theory based upon one underlying meaning relation. And then the two traditional conceptions of PRPs are synthesized into a theory of qualities and concepts. With these three unifications accomplished, it is possible to complete the final movement of the work—a solution to an array of outstanding problems from classical modern philosophy.

The first substantive chapter of the work, 'Intensionality', begins at the beginning: using a very small number of methodological assumptions, I retrace the intuitive motivation for the theory of PRPs. As I have said, one tenet of the work is that the theory of PRPs is part of logic. Specifically, it is part of natural logic, i.e., logic in the broad sense that includes the logic of natural language as a part. This tenet is defended in chapter 1, where I argue in particular that the best representation of intensionality in natural language is provided by a first-order intensional language that is just like a first-order extensional language except that it is fitted out with an intensional abstraction operation. This approach to intensional logic differs from the now prevalent one in that it locates the origin of intensionality in a single underlying intensional abstraction operation, rather than in an eclectic, open-ended list of operators such as modal operators, epistemic operators, deontic operators, etc.

In the next chapter, 'Intensional Logic', such an intensional language (called L_ω) is constructed. The semantics for L_ω, however, requires a new semantic method, one which harks back to the work of Boole, Peirce, and Schröder. This algebraic semantic method does not appeal to possible worlds even as a heuristic. The heuristic that is used is simply that of properties, relations, and propositions, taken at face value, and fundamental logical operations on properties, relations, and propositions. Using this new algebraic method, I define two notions of validity, one for the first traditional conception of intensional entities and one for the second traditional conception. Then, surprisingly as it might seem, the logics for L_ω relative to these two notions of validity are found to be both sound and complete. In this way I obtain two complete theories of PRPs, one ideally suited for modal matters and the other for intentional matters.

Chapter 3, the last chapter in part I, is devoted to the paradox of analysis, a particularly recalcitrant problem in the logic for intentional matters and one that has deep implications for philosophical method, philosophical psychology, and cognitive psychology. The problem may be put as follows: how, if correct, can a definition (or analysis) be informative? In recent years this important problem has been all but ignored. After a critical examination of Alonzo Church's resolution of the paradox, I show how to extend my new approach to intensional logic to include a more acceptable resolution.

Why is it that complete theories of PRPs are possible in the setting of first-order logic but not higher-order logic? The answer lies in the treatment of predication. The first-order approach, much like the approach taken by traditional logic in centuries past, treats the copula in natural language as a distinguished (2-place) logical predicate that expresses the fundamental logical relation of predication. The first-order logic for PRPs without the predication relation is complete. However, once the copula is singled out as a distinguished logical predicate, the logic for PRPs is rendered incomplete. (This can be proved by an application of Gödel's theorem that number theory is essentially incomplete; for once the logic for PRPs is equipped to represent the predication relation, it can then model number theory.) In this sense, then, it is not the infinite abstract ontology of logic—i.e., the ontology of PRPs—that is responsible for the incompleteness in logic; rather, it is a fundamental logical relation on that ontology—namely, the predication relation. This fact is hidden in the higher-order approach to logic since the notation for the predication relation is built into higher-order syntax right from the start. In spite of this, a popular thesis among modern logicians is that higher-order logic is a natural generalization of first-order logic. I take issue with this thesis in the first chapter of part II, 'Predication', where reasons are given for thinking that the first-order approach is the more natural and general of the two. It is in this context that the logical, semantical, and intentional paradoxes are considered. For not only is the predication relation responsible for the incompleteness in logic, but in addition it lies at the heart of these perplexing paradoxes. (Consider for example the analogue of Russell's paradox based on the predication relation.) Because of this, the logical, semantical, and intentional paradoxes can in the first-order setting be avoided, not by imposing restrictions on the existence of intensional entities, but rather by modifying what one would naively take to be the extension of the predication relation.

But what about set theory? Paradoxes arise there too. However, I argue in the next chapter, chapter 5, that set theory, unlike the theory of PRPs, is not rooted in natural logic and that it is instead born of certain confusions about natural logic. Of course, set theory, which is a relative newcomer on the intellectual scene, has proved to be very useful in both pure and applied mathematics. However, it is shown in this chapter that everything that set theory can do can be done equally well by the theory of PRPs. (Since this result holds for first-order pure set theory, which countenances sets of sets, sets of sets of sets, etc., it goes well beyond Russell's no-class construction, which works only for sets of non-sets.) The conclusion, then, is that there is no good logical or pragmatic reason for set theory. This shows that entities grounded in natural logic—namely, ordinary aggregates and their properties—may permanently take over the functions that were served on an *ad hoc* basis by the artificial abstract aggregates of set theory.

Now as far as the philosophy of mathematics is concerned, this is no mere changing of the guard. Since set theory is a poorly justified, artificial construct falling outside of logic proper, one may safely say that during its regime there has been no satisfactory philosophy of classical mathematics. However, since the theory of PRPs is a full-fledged part of natural logic, it can support a well justified logicist philosophy of mathematics. In chapter 6, 'Number', it is argued that this version of logicism is free from the standard criticisms of logicism, even those that proved fatal to the original logicism of Frege and Russell. This completes part II.

Up through this defense of logicism I assume a free and easy pragmatic posture toward three general issues—the relationship between intensional and extensional logic, the relationship between the semantic theory of truth and the theory of meaning, and the relationship between the two traditional conceptions of PRPs. Doing so considerably simplifies the investigation of several highly complex topics. However, these issues must be addressed before a fully unified theory can be attained. This is the first task of part III.

In chapter 1 it is concluded that intensionality in language can be traced to a single underlying intensional abstraction operation. For pragmatic reasons this operation is then treated as if it were a primitive, undefined operation. With matters left this way, one would conclude that there is a permanent bifurcation of logic into the extensional and the intensional. However, in chapter 7, 'Extensionality and Meaning', I show that it is formally possible to define the underlying intensional abstraction operation in a first-order extensional language for the theory of PRPs. This means that, if one wishes to, one may treat intensional sentences as mere syntactic transformations from fully extensional sentences. Doing this would enable one to conclude that intensionality in language is a mere surface phenomenon, that there is not really a bifurcation of logic into the extensional and the intensional, that logic is at bottom extensional logic. This conclusion, indeed, is just Carnap's thesis of extensionality.

In the study of the completeness and soundness problem for intensional logic it also proves convenient to characterize the semantics for intensional language by means of a theory of truth. But this approach to semantics altogether neglects the theory of meaning, as if the theory of truth were an autonomous domain. This situation also calls for remedy. Intuitively, a sentence is true if and only if what it means is true, i.e., if and only if it expresses a true proposition. At the same time, the concept of a true proposition is properly defined within the theory of PRPs quite independently of semantics (see §45). Therefore, the question of how to treat truth in language reduces to the question of how to characterize meaning. My discussion of this question centers around the classical Frege/Russell controversy about meaning.[7] Frege's theory is that there are two quite distinct meaning relations, expressing and naming. For Frege, every meaningful expression not containing free variables is a name, and every name expresses a sense (*Sinn*) and, in an ideal language at least, names a nominatum (*Bedeutung*). Russell's theory, on the other hand, is that there is only one fundamental meaning relation. Basically, naming is meaning restricted to names, and expressing is meaning restricted to predicates and sentences. For Russell, in contrast to Frege, predicates and sentences are not genuine names. Names do not express anything; they only name. And predicates and sentences do not name anything; they only express. Russell's theory certainly comes much closer to the commonsense theory of meaning than Frege's does, but is it as adequate? In chapter 7 I argue that it is. And this conclusion stands independently of the outcome of the Frege/Mill-Kripke controversy concerning the semantical properties of ordinary names in natural language. For this controversy deeply affects neither the Frege/Russell dispute over the general form a theory of meaning should take nor the general character the theory of PRPs should have.

Now predicates and open-sentences express properties and relations, and sentences express propositions. But which type of PRPs are these, which type of PRP is relevant to the theory of meaning? Do they conform to the first traditional conception or the second? Given Paul Grice's intentionalist analysis of meaning, it is easy to establish that the second traditional conception of PRPs is the one relevant to the theory of meaning, for it is this conception that is suited to intentional matters to begin with. But how, then, does the first traditional conception fit in? This question is taken up in chapter 8, 'Qualities and Concepts', whose content I will now describe in some detail.

Properties and relations play a primary role in the objective, non-arbitrary categorization and identification of objects, and in the description and explanation of change, and also in the constitution of experience. But not just any properties and relations can play these important roles; the ones that can are said to be *qualities* and *connections*. Among the myriad properties and relations, it is qualities and connections that determine the logical, causal, and phenomenal order of reality. Now when qualities and connections are combined by means of appropriate fundamental logical operations, sooner or later one comes to *conditions*. Conditions are the sort of things that can be said to obtain.

Intensional entities that are neither qualities, connections, nor conditions are ones that pertain primarily, not to the world, but instead to thinking taken in the broadest sense. Such PRPs are called *concepts* and *thoughts*. Consider the example of green and grue mentioned earlier. Whereas green is a genuine quality (specifically, a sensible quality), grue is only a concept (i.e., the concept expressed in English by the expression 'green if examined before *t* and blue otherwise'). As such, grue plays no primary role in the objective, non-arbitrary categorization and identification of objects; nor does it play a primary role in the description and explanation of change; nor does it play a primary role in the constitution of experience. Nevertheless, like other concepts, grue can play a role, even if a silly one, in thinking about the world. Now, from a purely logical point of view, the difference between qualities, connections, and conditions, on the one hand, and concepts and thoughts, on the other, is that the former conform to the first traditional conception of intensional entities; that is, qualities, connections, and conditions are identical if and only if they are necessarily equivalent. However, though necessary equivalence is a necessary condition for the identity of concepts and thoughts, it is not a sufficient condition. For concepts and thoughts conform to the second traditional conception; they must have unique, non-circular definitions (analyses). In this way, then, qualities, connections, and conditions pertain more to modal matters (necessity, possibility, etc.), and concepts and thoughts pertain more to matters of intentionality.

Just as conditions are the sort of things that can be said to obtain, so thoughts are the sort of things that can be said to be true. According to the commonsense theory of truth, a thought is true if and only if the condition to which it corresponds obtains. But what is this relation of correspondence that holds between thoughts and conditions? Concerning this question modern philosophy from Descartes and Locke onward has been dominated almost exclusively by doctrines of representationalism. However, representationalism leaves the relation of correspondence veiled in mystery and metaphor: at most, thoughts and concepts give way to other thoughts and concepts, *ad infinitum*; one never gets to the real thing. The alternative to representationalism is realism, a doctrine whose origins may be traced back to certain works by Plato and Aristotle. According to realism, we are not forever caught in a net of representations. On the contrary, when a thought or concept is fully and properly analysed, we eventually come, not to still further representations, but instead to the kind of properties and relations that actually give the world and our experience its structure; that is, we eventually come to genuine qualities and connections. To put the point the other way around, according to realism, there are certain fundamental logical operations such that, when qualities and connections are combined by means of them, what we get are concepts and thoughts; all concepts and thoughts are obtained from qualities and connections (plus perhaps subjects of singular predications) by means of these fundamental thought-building operations. Now consider a given thought, and consider the procedure by which this thought is built up from genuine qualities and connections by means of fundamental thought-building operations. Next suppose that the same qualities and connections are combined using the very same procedure except that this time the use of fundamental thought-building operations is replaced by the use of the fundamental condition-building operations. The result is, of course, a condition. But what condition is it? The answer seems plain: it is none other than the condition to which the original thought *corresponds*. And in this way, realism attains something that seems forever out of reach of representationalism, namely, a purely logical analysis of the relation of correspondence.

This realistic synthesis of the two traditional conceptions of PRPs is formalized at the close of chapter 8. Then in the final two chapters, 'Logic' and 'Mind', the resulting theory of qualities and concepts is used to obtain solutions to a number of outstanding problems in classical modern philosophy, problems that have resisted solution to a large extent because they are typically thought of in terms of representationalism rather than realism. The first three problems concern the definitions of truth, necessity, and analyticity. Given the realistic analysis of the correspondence relation and given the fact that qualities, connections, and conditions conform to the first traditional conception of PRPs (i.e., they are identical if and only if they are necessarily equivalent), it is possible to give purely logical definitions of what it is for a thought to be true, necessary, analytic, or valid. Furthermore, it is possible to give a purely logical definition of what it is to be a necessary connection, something Hume vigorously denied could be done. And given this, we are perhaps a step closer to a purely logical analysis of causation.

The final three problems to which the theory of qualities and concepts is applied are intentionality, mind, and consciousness, the very hallmarks of modern philosophy. According to Franz Brentano, an intentional phenomenon is one that is *about* something else, even if that something else does not exist. Notice, however, that thoughts and concepts are exactly the sort of thing that can be about something else, even if that something else does not exist. Beginning with this insight, I define an intentional connection as one that can hold in a certain way between an individual particular and a thought or a concept. And generalizing on this definition, I define a mental connection—intentional or experiential—as one that can hold in a certain way between an individual particular and a thought or concept or a quality, connection, or condition. According to this definition, then, mental connections are different from physical connections, for the latter can hold only between particulars and particulars, or perhaps between particulars and locations, particulars and times, particulars and stuffs, etc. The intuition behind this definition is that thoughts and concepts play a special role in thinking, and qualities, connections, and conditions play a special role in experience. And none of them plays any such role in exclusively physical processes. Those who hold otherwise would seem to have forgotten a category distinction between physical and mental connections.

Now suppose that one day we should design and build a machine that performs physically as we do, behaviorally and mechanically. A natural question to ask is whether the machine has a mind. According to behaviorism and materialistic versions of functionalism, this question is identical to the question of whether the machine behaves or functions physically as we do. But *ex hypothesi* we know that it does; *that* is not our question. We want to know something else, but what is it? What we want to know is whether the machine actually functions mentally. But this is to say, we want to know whether it stands in genuine mental connections to things. For intuitively, a thing functions mentally if and only if it really does stand in mental connections to things. (Connections, it will be recalled, are the fundamental kind of relations which, together with genuine qualities, serve to fix the logical, causal, and phenomenal order of reality.) It is not enough that the machine should merely behave or function *as if* it is mentally connected to things. This difference is what makes all the difference. Yet by using the theory of qualities and concepts we are able to give a purely logical definition of what a genuine mental connection is. And in fact this definition can be used to give a purely logical definition of consciousness itself.

The overall movement in the work is thus toward a unified logical realism. Solutions to fundamental metaphysical problems are found in neither empiricism nor naturalism nor idealism but rather in logic, which underlies the very exercise of reason.

4. Critical Survey of Alternate Approaches*

In developing this theory of qualities and concepts I have been guided by several desiderata, which I will now simply state.

Desiderata

(a) It is desirable that the theory should provide at least a framework for solving the following family of classical puzzles in the philosophy of logic and language:

Classical Puzzles

(1) substitutivity failures involving co-extensive expressions in modal and intentional contexts (§§8, 39)
(2) substitutivity failures involving necessarily equivalent expressions in intentional contexts (§16)
(3) the paradox of analysis (§§18–20)
(4) Mates' puzzle concerning substitutivity failures for synonyms in intentional contexts (§§18, 39)
(5) quantifying-in, i.e., the external quantifiability of certain singular terms in modal and intentional contexts (§§7, 11)
(6) anomalies involving indexicals in intentional contexts (§39)
(7) Geach's puzzle about intentional identity (§39)

* Knowledge of the material surveyed in this section is not required for an understanding of the rest of the book. Discussions of the various desiderata may be found in the sections indicated.

(8) the logical paradoxes, e.g., the property-theoretic analogue of Russell's paradox (§26)
(9) the semantical and intentional paradoxes, e.g., the semantical and intentional versions of the liar's paradox (§26)
(10) Frege's puzzle, i.e., how can '$a = b$' be true yet different in meaning from '$a = a$'? (§38).

(b) It is desirable that the theory should in its formal statement constitute an idealized representation of natural language having the following programmatic features:

Programmatic Features

(11) it has sound and complete logics for modal matters and for intentional matters (§§15–17)
(12) it passes the Langford-Church translation test (§8)
(13) it satisfies Davidson's finite-learnability requirement, i.e., it has a finite number of undefined constants (§§8, 12, 37)
(14) it has no *ad hoc* existence restrictions imposed by stratification according to logical type (§§10, 22)
(15) it makes no ontological commitment to non-actual possibilia (§§13, 39, 46)
(16) it represents 'believes' as a 2-place predicate, 'is true' and 'is necessary' as 1-place predicates, and 'that'-clauses as singular terms (§6)
(17) it is syntactically first-order (§§10, 21–6)
(18) it has a Russellian semantics, i.e., its semantics specifies only what the genuine names name and what the predicates and sentences express (§38)
(19) it is consistent in its semantics with Mill's theory of ordinary names (§§38–9)
(20) it is consistent with Carnap's thesis of extensionality (§37).

(c) It is desirable that the theory should yield the following applications:

Applications

(21) an analysis of number (§32)
(22) a definition of truth for propositions (§45)
(23) a definition of necessity (§46)

(24) a definition of analyticity (§47)
(25) an analysis of intentionality (§48).

It is, of course, understood that these analyses and definitions are to be non-circular.

Some of the desiderata might appear tendentious. Each one, however, will be discussed in some detail elsewhere in the book, and I hope that such discussion will help to resolve objections.

Before I launch into the work proper it should be helpful to have an overview of the various leading approaches to our general subject area and the success that these approaches have in satisfying the desiderata. The most efficient way to provide such an overview is in chart form. Of necessity, such a chart will be provisional in character: the material to be represented is very complicated and many of the entries will obviously be arguable. Nevertheless, since a thorough discussion of past approaches would easily comprise a book of its own, and since my purpose here is to advance a new approach and not to review the merits and defects of past ones, the use of a provisional chart would seem to be the best way to meet the present need. The chart is intended only as a tool to be used in obtaining an overview of a complicated family of problems and attempted solutions.

The following are the approaches represented on the chart; salient features of the approaches are mentioned where that is necessary in order to distinguish one approach from another:

Approaches

(1) the theory of qualities and concepts
(2) Frege's approach (*Die Grundlagen der Arithmetik*, 'Funktion und Begriff', 'Über Sinn und Bedeutung', *Grundgesetze der Arithmetik*)
(3) Church's Alternative (2)—all and only necessarily equivalent expressions are counted as synonymous ('A Formulation of the Logic of Sense and Denotation', 'Outline of a Revised Formulation of the Logic of Sense and Denotation')
(4) Church's Alternative (0)—all and only synonymous isomorphic expressions are counted as synonymous ('A Formulation', 'Outline', 'Intensional Isomorphism and Identity of Belief')

(5) Russell's approach ('On Denoting', 'Mathematical Logic as Based on the Theory of Types', *Principia Mathematica*, 'The Philosophy of Logical Atomism')

(6) Carnap's approach—'is necessary' is treated as an operator, not a predicate; in both the object language and the metalanguage quantification over non-actual possibilia is avoided in favor of the wholly syntactical state-descriptions approach (*Meaning and Necessity*)

(7) Hintikka's approach—'believes' and 'is necessary' are both treated as operators, not predicates; *à la* Davidson, the theory of meaning is equated with the theory of truth; possible-worlds semantics is used, but explicit quantification over non-actual possibilia is avoided in the object language (*Models for Modalities*)

(8) Montague's approach—'believes' and 'is necessary' are both treated as predicates; possible-worlds semantics is used, but explicit quantification over non-actual possibilia is avoided in the object language ('Pragmatics', 'Universal Grammar', 'The Proper Treatment of Quantification in English'; it will be assumed here that the axiomatic theories given by Daniel Gallin in *Intensional and Higher-Order Modal Logic* are part of Montague's approach)

(9) David Lewis' approach—propositions are explicitly treated in the object language as functions from possible worlds onto truth values; properties, with functions from possible worlds onto sets of possible individuals; etc.; it will be assumed here that Lewis wishes to quantify explicitly over non-actual possibilia in the object language ('General Semantics', *Counterfactuals*, etc.)

(10) Scheffler's inscriptional approach—$\ulcorner x$ says that $A\urcorner$ is analysed as $\ulcorner(\exists y)(y$ is-an-A-inscription & x utters $y)\urcorner$, where $\ulcorner$is-an-A-inscription$\urcorner$ is a primitive 1-place predicate that is satisfied by all and only inscriptions synonymous to A ('An Inscriptional Approach to Indirect Quotation')

(11) Davidson's approach—$\ulcorner x$ says that $A\urcorner$ is parsed as $\ulcorner x$ says that: $A\urcorner$, where 'that' is a demonstrative referring to what follows it; the theory of meaning is assimilated to the theory of truth ('On Saying That', 'Truth and Meaning', etc.)

(12) Quine's syntactical approach—$\ulcorner x$ believes that $A\urcorner$ is analysed as $\ulcorner x$ believes-true $\ulcorner A\urcorner\urcorner$ (§44 *Word and Object*)

(13) Quine's primitive-predicate approach—⌜x believes that A⌝ is analysed as ⌜x believes-that-A⌝, where the latter is a primitive 1-place predicate that is satisfied by all and only things who believe that A (§44 *Word and Object*).

I will include under (12) and (13) Quine's familiar approaches to definite descriptions, names, number, and truth (*Word and Object*, *Mathematical Logic*, 'On What There Is', 'Notes on the Theory of Reference', etc.).

In evaluating these approaches relative to the desiderata, I will use the following five grades:

Yes No ? − +

To see the force of these grades, suppose that a desideratum D concerns some classical puzzle and that an approach A explicitly contains a candidate solution to the puzzle. Then, if successful, A receives the grade 'Yes'; if unsuccessful, A receives the grade 'No', and if it is uncertain whether the solution is successful, A receives the grade '?'. On the other hand, suppose that A does not explicitly contain a solution to the puzzle. Can A be supplemented with a successful explicit solution to the puzzle? If there seems to be a barrier to doing this, then A will receive the grade '−'. If there does not, then A will receive the grade '+'. The '+' grade does not imply that a barrier does not exist; it only means that one is not in evidence. In the event that desideratum D concerns, not some classical puzzle, but rather one of the programmatic constraints on A or one of the applications of A, the five grades are used analogously. Thus, we have the following:

Key

Yes: explicit stance taken—outcome successful
No: explicit stance taken—outcome unsuccessful
?: explicit stance taken—outcome uncertain
−: no explicit stance taken—evident barrier to success
+: no explicit stance taken—no evident barrier to success.

In grading the various theories I will make use of another convention. Many of the approaches presuppose the doctrine that ordinary names are synonymous with definite descriptions. This descriptivist doctrine, of course, contradicts Mill's theory of ordinary

DESIDERATA	APPROACHES: 1. Quality and concept	2. Frege	3. Church: alt. 2	4. Church: alt. 0	5. Russell	6. Carnap	7. Hintikka	8. Montague	9. D. Lewis	10. Scheffler	11. Davidson	12. Quine: syntactical	13. Quine: prim. pred.
1. Substitutivity failures: co-extensives	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2. Substitutivity failures: nec. equivalents	Yes	Yes	No	Yes	Yes	No	No	No	No	Yes	Yes	Yes	Yes
3. Paradox of analysis	Yes	–	–	No	–	No	–	–	–	+	+	–	–
4. Mates' puzzle	Yes	+	+	?	+	+	+	+	+	+	+	+	+
5. Quantifying-in	Yes	–	No	No	Yes	–	Yes	Yes	Yes	+	–	No	–
6. Indexicals in intentional contexts	Yes	No	–	–	?	+	+	–	–	+	+	+	–
7. Intentional identity	Yes	+	+	+	+	+	+	?	+	+	+	+	+
8. Logical paradoxes	Yes	No	Yes	Yes	Yes	Yes	+	Yes	Yes	+	+	Yes	Yes
9. Semantical and intentional paradoxes	Yes	+	+	Yes	Yes	+	+	+	+	+	+	+	+
10. Frege's puzzle	Yes	Yes	Yes	Yes	Yes	Yes	–	Yes	Yes	+	No	–	–
11. Sound and complete logics for modalities and intentional matters	Yes	No	No	No	No	–	+	No	–	–	?	–	–
12. Langford-Church translation test	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No	No	Yes
13. Davidson's learnability requirement	Yes	No	No	No	No	Yes	Yes	–	–	No	Yes	Yes	No
14. Not a type-theory	Yes	No	No	No	No	No	Yes	No	No	Yes	Yes	Yes	Yes
15. No commitment to non-actual possibilia	Yes	Yes	Yes	Yes	Yes	Yes	?	?	No	Yes	Yes	Yes	Yes
16. 'Necessary', 1-ary predicate; 'believe', 2-ary predicate; 'that'-clause, sing. term	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	+	No	No	No
17. Syntactically first-order	Yes	No	No	No	No	No	Yes	No	No	Yes	Yes	Yes	Yes
18. Russellian semantics	Yes	No	No	No	?	No	No	No	No	+	No	No	No
19. Consistent with Mill's theory of names	Yes	No	No	No	No	No	+	No	No	+	Yes	No	No
20. Thesis of extensionality	Yes	Yes	Yes	Yes	No	No	No	No	?	Yes	Yes	Yes	Yes
21. Definition of number	Yes	Yes	Yes	Yes	Yes	+	+	+	+	+	+	Yes	Yes
22. Definition of truth for propositions	Yes	+	Yes	Yes	+	–	–	+	+	–	–	–	–
23. Definition of necessity	Yes	–	Yes	–	–	No	–	–	–	–	–	–	–
24. Definition of analyticity	Yes	–	–	–	–	No	–	–	No	–	–	–	–
25. Analysis of intentionality	Yes	–	–	–	–	–	–	–	–	–	–	–	–

names. In grading those approaches that presuppose the descriptivist doctrine, I will proceed as if this doctrine were correct, but I will give a 'No' grade on desideratum 19, consistency with Mill's theory. Incidentally, this sort of interdependence, which occurs at numerous points in the chart, is one reason for the chart's provisional character.[8] Finally, the grades given to the theory of qualities and concepts are surely tentative since, unlike the other approaches, this one has not yet been subjected to critical scrutiny.

As I have indicated, I feel uncertain about several of the grades. However, a number of the more controversial ones will be discussed in succeeding chapters. Readers with misgivings might find some satisfaction there.

5. Epistemological Note

I wish to say a word about those epistemological attacks on theories of PRPs made from the point of view of naturalistic empiricism. Although this is not the place to attempt a thorough rebuttal, I will sketch the general line of defense that I am inclined to take.

I believe that the theory of qualities and concepts has a traditional *a priori* justification: the theory is part of logic, and logic is knowable *a priori*. I believe that the theory in addition has a traditional *a posteriori* justification: at least some of one's own conscious states—uninterpreted experiental states or conscious intentional states—are directly evident, and the theory of qualities and concepts constitutes the best formulation of the logic for these directly evident matters. Yet unless worked out in detail, these two replies are unlikely to persuade philosophers with deep naturalistic empiricist convictions. This is no worry, however, for the empiricist attack can be answered within its own ground rules without invoking a competing rationalist or Cartesian-foundationalist theory of knowledge.

The naturalistic empiricist holds that a theory is justified if it belongs to our best composite theory of the world. However, if our best composite theory includes a mentalistic psychology and an intensional theory of meaning, then it will include the theory of qualities and concepts. The reason for this is twofold. First, as before, the theory of qualities and concepts constitutes the best formulation of the logic for psychological matters; secondly, it provides the best background theory for constructing theories of meaning. Up to now a mentalistic psychology and an intensional theory of meaning have always found their way into our best composite theory. To think that this situation will ever change looks more like an article of faith than like a rational belief. Therefore, unless unreasonable double standards are invoked, it appears that such theories—and, in turn, the theory of qualities and concepts—are justified from the naturalistic empiricist point of view.

In closing I should like to turn the tables on naturalistic empiricism. The very vocabulary in which its doctrines are stated—'justification', 'best theory', etc.—simply cannot be explained satisfactorily outside a logical framework like that provided by the theory of qualities and concepts. Thus, if naturalistic empiricism is wedded to the attack on theories of PRPs, epistemologically it is radically self-defeating. Epistemologically self-approving theories may all be expected to use a logical framework like that provided by the theory of qualities and concepts.[9]

The theory of qualities and concepts achieves a level of conceptual clarity uncommon in the special sciences endorsed by naturalistic empiricists. According to naturalistic empiricism even the most basic logical concepts must be explained in naturalist and empiricist terms. However, the natural order of explanation is certainly in the opposite direction at least for the most general naturalistic concepts such as cause, matter, mind, species, nature, etc. These concepts will remain obscure until they are explained in terms of still more fundamental logical concepts, concepts used simply in the exercise of reason.

Part I

A Complete Foundation

1

Intensionality

Intensionality in logic and language is a phenomenon that has been recognized for over two millennia, and still there is no adequate theory for it. My investigations of intensionality will begin with an elementary inquiry into the origins of intensionality in natural language. Some surprisingly simple arguments will expose defects in what is today the leading treatment of intensionality, the multiple-operator approach. The best representation of intensionality, it will turn out, is one that explicitly appeals to properties, relations, and propositions. In this, the theory of PRPs is seen to be undeniably part of logic.

6. Intensional Abstraction

Consider the following intuitively valid argument:

(I) Whatever x believes is necessary.
Whatever is necessary is true.

∴ Whatever x believes is true.

Suppose that 'is necessary' and 'is true' are treated as 1-place predicates and 'believes', as a 2-place predicate.* Then, the above argument can be represented as valid in any standard quantifier logic:

$$\begin{array}{ll} (\mathrm{I}') & (\forall y)(xB^2y \supset N^1y) \\ & (\forall y)(N^1y \supset T^1y) \\ \hline & \therefore\ (\forall y)(xB^2y \supset T^1y). \end{array}$$

Now in theoretical matters, if a currently accepted theory can be easily and naturally employed to account for new data, then other

* In this work when I mention linguistic expressions I will usually follow the convenient convention of autonymous use, by which a simple expression names itself and a concatenation of simple expressions names their concatenation. But where clarity demands, I will shift to the use of single quotes; when I do this, I will sometimes take the liberty to use them for the kind of variable quotation achieved more properly by Quinean corner quotes. I reserve double quotes for use as scare quotes.

things being equal it is desirable to do so. In the science of logic, the currently accepted theory includes quantifier logic. By treating 'is necessary' and 'is true' as 1-place predicates and 'believes' as a 2-place predicate, we can easily and naturally account for the validity of (I) in a currently accepted theory, namely, quantifier logic. Other things being equal, it is therefore desirable to do so.[1]

Now consider another intuitively valid argument, where A is any formula:

(II) Whatever x believes is true.
x believes that A.
———————————————
∴ It is true that A.

Suppose, as was just suggested, that we do treat 'is true' as a 1-place predicate and 'believes' as a 2-place predicate. In this case, we seem to be left with no alternative but to parse the second and third lines of (II) as follows:

$$\underbrace{x}\ \underbrace{\text{believes}}\ \underbrace{\text{that } A}$$
$$\underbrace{\text{It is true}}\ \underbrace{\text{that } A}$$

where 'that A' is counted as a singular term syntactically. I do not wish to beg any questions here about the philosophical treatment of 'that'-clauses. For this reason, I will introduce a philosophically neutral notation. I have in mind the bracket notation introduced by Quine for somewhat similar purposes (§35 *Word and Object*). For the moment I leave open what semantical significance the bracket notation will have, and the possibility of indirectly defining the bracket notion will also be left open here.[2] When this bracket notation is adopted, (II) can be naturally represented as follows:

(II′) $(\forall y)(xB^2y \supset T^1y)$
$xB^2[A]$
———————————————
$\therefore\ T^1[A]$.

The conclusion of (II′) is straightforwardly derivable from the two premises by an application of universal instantiation (UI) and modus ponens (MP), two rules of inference valid in standard quantifier logic. Thus, one can bring argument (II) within the scope of standard quantifier logic simply by adopting the hypothesis that 'that'-clauses are singular terms representable with the bracket notation. To successfully represent (II), one needs no new logical principles, and one needs no knowledge about the logic of expressions occurring within [A]. It would seem, therefore, that relative to the framework of quantifier logic, (II′) is the simplest way to represent (II). Thus, on the assumption that the logic for the new singular terms [A] can be satisfactorily worked out, I conclude that it is desirable to treat 'that'-clauses in natural language as singular terms that may be represented by means of the bracket notation.[3]

Summing up, I conclude that it is desirable to treat 'is true' and 'is necessary' as 1-place predicates, 'believes' as a 2-place predicate, and 'that'-clauses as (defined or undefined) singular terms. (This conclusion is just desideratum 16 from §4.)

7. Quantifying-in

Consider the following argument:

(III) x believes that he believes something.
____
∴ There is someone v such that x believes that v believes something.

There is a reading according to which (III) is intuitively valid. This reading provides an example of the logical phenomenon of quantifying-in. It is desirable that all valid cases of quantifying-in should be representable by an ideal logical theory. (This is desideratum 5 from §4.)

Putting desiderata 5 and 16 together, one obtains an important derived desideratum. Consider the following instance of argument (II):

(IV) Whatever x believes is true.
x believes that v believes something.
____
∴ It is true that v believes something.

In view of desiderata 5 and 16 it is desirable to represent (IV) as follows:

(IV′) $(\forall y)(xB^2y \supset T^1y)$
$xB^2[(\exists u)vB^2u]$
____
$\therefore\ T^1[(\exists u)vB^2u]$.

I conclude by analogy that it is desirable to represent (III) in the following way:

$$\text{(III}')\qquad \frac{xB^2[(\exists u)xB^2u]}{\therefore\ (\exists v)xB^2[(\exists u)vB^2u].}$$

What is important about this is that the occurrence of v in the singular term $[(\exists u)vB^2u]$ is an *externally quantifiable* occurrence of a variable.[4] I am thus led to conclude that 'that'-clauses ought to be treated as singular terms which may contain externally quantifiable occurrences of variables.

It will be convenient to represent in some perspicuous way which variables within $[A]$ are externally quantifiable. Let δ be the sequence of externally quantifiable variables in $[A]$. Then, I will rewrite $[A]$ as $[A]^\delta$. So, for example, I will rewrite (III′) as follows:

$$\frac{xB^2[(\exists u)xB^2u]^x}{\therefore\ (\exists v)xB^2[(\exists u)vB^2u]^v.}$$

This allows the externally quantifiable variables to be spotted at a glance.

8. Informal Interpretation

I have concluded that it is desirable to represent 'that'-clauses with the bracket notation. Up to now I have left open how this bracket notation should be interpreted and, in particular, what sort of entity corresponds semantically to a given singular term $[A]$. In order to answer this question we must consider desideratum 1 from §4, which concerns *prima facie* failures of substitutivity of co-extensive expressions within 'that'-clauses.

Consider the following argument:

(V) x believes that everything runs.
Everything runs if and only if everything walks.
———————————————————————
∴ x believes that everything walks.

Argument (V) is *prima facie* an instance of the principle of the substitutivity of materially equivalent formulas. However, (V) is intuitively invalid. Thus, it constitutes a *prima facie* violation of the principle of the substitutivity of materially equivalent formulas.

Now consider the following related argument:

(VI) x wonders whether y is the author of *Waverley*.
$y =$ the author of *Waverley*.
$\therefore$ x wonders whether $y = y$.

(VI) is a *prima facie* instance of the principle of the substitutivity of co-referential singular terms. However, there is a reading of (VI) according to which it too is invalid. Thus, we have a *prima facie* violation of this substitutivity principle. Desideratum 1 is simply that arguments containing *prima facie* violations of these two substitutivity principles ought to be represented as invalid in an adequate logical theory.

In the bracket notation (V) would be represented as follows:

(V′)
$$\begin{array}{l} xB^2[(\forall y)Ry] \\ (\forall y)Ry \equiv (\forall y)Wy \\ \hline \therefore\ xB^2[(\forall y)Wy]. \end{array}$$

And the invalid reading of (VI) would be represented as follows:

(VI′)
$$\begin{array}{l} xW^2[y = (\iota z)(Az)]^y \\ y = (\iota z)(Az) \\ \hline \therefore\ xW^2[y = y]^y \end{array}$$

where in the first premise the definite description $(\iota z)(Az)$ has narrow scope. Now in order for (V′) and (VI′) to qualify as invalid arguments, what sort of entities must correspond semantically to the singular terms $[(\forall y)Ry]$, $[(\forall y)Wy]$, $[y = (\iota z)(Az)]^y$, and $[y = y]^y$? Both here and in what follows my intention is to use the notion of semantical correspondence in as neutral a way as possible. By doing so, I wish to avoid committing myself to any particular semantical method. And also I wish to take into account the fact that 'that'-clauses might be contextually defined singular terms and, hence, that they might bear no simple semantical relation (e.g., the naming relation) to anything. Even if 'that'-clauses are contextually defined singular terms, their use nevertheless produces ontological commitments;[5] thus, in asking what sort of entity corresponds semantically to the singular terms $[A]$, we are at the very least asking to what sort of entity the use of 'that'-clauses ontologically commits us.

The nominalistic answer to the above question is that linguistic entities—either formulas or inscriptions of formulas—are what correspond semantically to 'that'-clauses. Generally speaking, there are two methods by which one can develop the nominalistic answer in detail. According to the first method, a formula such as $xB^2[(\forall y)Ry]$ is treated in such a way that it contains—at least when fully analysed—either a name for, or a structural description of, a particular formula or inscription. On the second method, a formula such as $xB^2[(\forall y)Ry]$ is treated in such a way that even when fully analysed it does not contain any such metalinguistic name or structural description. Carnap's approach and Quine's syntactical approach are instances of the first method. Scheffler's approach is an instance of the second method.

A fatal difficulty in the first method is that it leads to violations of desideratum 12, the Langford-Church translation test. The argument that these nominalistic analyses lead to faulty translation is familiar enough that I will not go over it here.[6]

The nominalist's second method, by contrast, does satisfy desideratum 12. However, it evidently must do so at the price of violating desideratum 13, Davidson's learnability requirement. (Davidson's learnability requirement is that an idealized representation of natural language should have a finite number of undefined primitive constants.) To see why this is so, let us consider Scheffler's approach as an example. According to this approach, a singular term $[A]$ would be contextually analysed as follows:

$$\ldots [A] \ldots \mathit{iff}_{\mathrm{df}}\ (\exists v_k)\ (v_k \text{ is-an-}A\text{-inscription } \& \ldots v_k \ldots)$$

where 'is-an-A-inscription' is an undefined primitive predicate that is satisfied by all and only inscriptions synonymous to A. However, since there are an infinite number of distinct 'that'-clauses in natural language, there must be an infinite number of distinct singular terms $[A]$. Therefore, Scheffler's approach requires an infinite number of undefined primitive predicates 'is-an-A-inscription'. The fact that Scheffler's approach requires an infinite number of undefined primitive predicates not only blocks learnability but also blocks the systematization of the internal logic of 'that'-clauses.

The above considerations, together with a number of others,[7] lead me to conclude that linguistic entities, whether formulas or inscriptions of formulas, are not the sort of entity that correspond semantically to the singular terms $[A]$. And the same conclusion goes for sequences or sets of linguistic entities, or indeed any other kind of object that is linguistic in character.

So what sort of entities do correspond semantically to the singular terms $[A]$? Whatever they are they must render arguments such as (V′) and (VI′) invalid. Further, they must lead to no violations of the Langford-Church translation test. And finally, they must make possible the kind of finitistic treatment of language called for by Davidson's learnability requirement. Now we shall see that these features are had by propositions, which are one kind of intensional entity. (Intensional entities are ones that need not be identical even if they are identical in extension.) To be sure, these features are also had by certain other entities that are not in themselves intensional. But upon analysis these other entities seem always to involve some sort of appeal to intensional entities. (For example, these alternate objects might be sets—or sequences or mereological sums—of intensional entities.) So of the choices available, propositions taken on their own make for the most natural answer to the question. Therefore, other things being equal one may conclude that propositions should be identified as the semantical correlata of the singular terms $[A]$.

Why do propositions meet our needs? Why, for example, does argument (V′) come out as invalid when propositions are identified as the semantical correlata of the singular terms $[(\forall y)Ry]$ and $[(\forall y)Wy]$? The answer is simply that the propositions semantically correlated with these two singular terms are not the same. And this is so even though these propositions have the same extension, i.e., even though they have the same truth value. And why do propositions make it possible to pass the Langford-Church translation test? The answer is that propositions are extralinguistic entities. And thus, when 'that'-clauses are given the recommended interpretation, they can be translated into other languages independently of problematic names for, or structural descriptions of, linguistic entities. Finally, how do propositions make it possible to meet Davidson's learnability requirement? The answer to this question is by no means obvious. Indeed, all previous theories of propositions have failed on this score. (See desideratum 13 on the chart.) This is one of the outstanding problems that a new theory of PRPs must surmount. But it turns out that the syntactic and semantic construction in §§12–14 solves it.

9. The Origin of Intensionality in Language

Intensional entities are, as I have said, entities that can be different from one another even though they are identical in extension. Propositions are 0-ary intensional entities; properties, 1-ary intensional entities; and relations, n-ary intensional entities, for $n \geq 2$. In view of this, there is a natural generalization of the bracket notation which provides singular terms for intensional entities of any finite degree. Let A be any well-formed formula, and let $v_1, \ldots, v_m$ be distinct variables, where $m \geq 0$. (I permit there to be free variables in A that are not among these variables $v_1, \ldots, v_m$.) Then, $[A]_{v_1 \ldots v_m}$ is a singular term whose semantical correlate is an intensional entity of degree m. If $m = 0$, the semantical correlate of this singular term is the proposition that A; if $m = 1$, the semantical correlate is the property of being something v_1 such that A; if $m > 1$, then the semantical correlate is the relation among $v_1, \ldots, v_m$ such that A.

In §6 and §7 it was argued from the point of view of logical syntax that certain complex nominative expressions in natural languages—namely, 'that'-clauses—are best represented by singular terms of the sort provided by the bracket notation $[A]_{v_1 \ldots v_m}$, where $m = 0$. There are analogous arguments to show that certain other complex nominative expressions in natural language—namely, gerundive and infinitive phrases—are best represented as singular terms of the sort provided by our generalized bracket notation $[A]_{v_1 \ldots v_m}$, where $m \geq 1$. By this route, then, the theory of PRPs is found to be part of the logic for natural language.

What is logically distinctive about these singular terms $[A]_\alpha$ is that expressions occurring within them do not obey the substitutivity principles of extensional logic. Thus, when a formula A is enclosed within square brackets (followed by appropriate subscripts), an intensional context is generated. This bracketing operation may therefore be viewed as a generalized *intensional abstraction operation*.

According to the now dominant tradition of C. I. Lewis, Carnap, Hintikka, Kripke, *et al.*, intensionality in natural language originates with a diverse, open-ended list of primitive operators, including, e.g., a strict-implication operator, modal operators, deontic operators, epistemic operators, an assertion operator, a causal-explanation operator, a would-be-fact operator, probability operators, ... The intensional abstraction approach to intensionality is different. Suppose that a multiple-operator theorist has the need for a primitive n-place intensional operator $\mathcal{O}^n$. In this case, the advocate of intensional abstraction will instead have an associated n-place primitive predicate O^n. Thus, where the operator theorist has a new category of operator sentences $\mathcal{O}^n(A_1, \ldots, A_n)$, I will simply have the atomic sentences $O^n([A_1], \ldots, [A_n])$. (As I will show, it is as easy to state the semantics for O^n and $[A_i]$ as it is to state the semantics for $\mathcal{O}^n$.) The intensional-abstraction approach has two distinct advantages over the multiple-operator approach. The first has already been discussed: since these diverse primitive operators cannot take singular terms as arguments, there are infinitely many intuitively valid arguments that cannot be represented by this approach. (Argument (I) in §6 is one such argument.) The second advantage is that the general theory of intensionality that emerges on the operator approach is eclectic and incomplete at best. On the intensional-abstraction approach, however, there is a simple and general theory of intensionality: all intensionality in natural language (or at least all intensionality treatable by some operator or other) has a single origin, namely, a generalized intensional abstraction operator. Because of the simplicity and generality of this approach, I conclude that it provides the best provisional representation of intensionality in natural language.[8]

10. First-Order Language

Early in §6 I asserted that in the science of logic the currently accepted theory includes quantifier logic. At that time I chose to defer the question of whether we ought to adopt a first-order or a higher-order formulation as our standard quantifier logic. I will now take up this question. In chapter 4 I will defend the position that the first-order approach is the more natural and general of the two. Here, I will simply list my reasons for thinking that, from the point of view of formal logico-linguistic theory, the first-order approach is superior to the higher-order approach.

First, first-order quantifier logic is complete; higher-order quantifier logic is not. At the same time, the consistency of first-order quantifier logic is less open to doubt than that of the higher-order counterpart. Other things being equal, it is desirable to construct a new theory within a theoretical framework that is complete and whose consistency is as little open to doubt as possible. Thus, other things being equal, it is desirable to construct a new theory within the framework of a first-order quantifier logic as opposed to higher-order quantifier logic.[9]

Secondly, if the first-order approach to quantifier logic is taken, then, as I will show, it is possible to construct a sound and complete logic for the intensional abstraction operation and, hence, for modal matters and for intentional matters (see desideratum 11). Such a result is out of reach if the higher-order approach is taken.

Thirdly, when the higher-order approach is taken, linguistic predicates and sentences (open or closed) are treated as linguistic subjects. This would seem to open up the possibility of new instances of Frege's '$a = a$'/'$a = b$' puzzle. (This possibility is what lies behind Church's worry about the adequacy of a Russellian semantics to characterize the semantics for *Principia Mathematica*. See §23 and §38.) Thus, in the case of higher-order languages Russellian semantics is problematic; not so in the case of first-order languages. (See desideratum 17.)

Fourthly, in order to avoid the logical and intentional paradoxes (see desiderata 8 and 9), the higher-order approach usually incorporates infinitely many distinct sorts of variables which carry with them an implicit commitment to a theory of logical types. (See desideratum 14.) Type theory, however, imposes especially implausible existence restrictions on PRPs, restrictions that in most cases play no direct role in the avoidance of the paradoxes. The first-order approach, by contrast, can easily avoid the logical and intentional paradoxes without appealing to type theory.

Fifthly, suppose that, in order to avoid the logical and intentional paradoxes, a given higher-order theory incorporates infinitely many distinct sorts of variables. In this case, it will be forced into a violation of desideratum 13, Davidson's learnability requirement. To see this, consider any "transcendental" predicate in natural language, i.e., any predicate in natural language whose extension cuts freely across presumed type boundaries. The 2-place predicate 'contemplate' is an example of such a predicate. Since the open sentence '$(\exists x)x$ contemplates y' is satisfiable by objects in every logical type, 'contemplate' would have infinitely many primitive counterparts $C^2_{o\alpha}$, one for each sort of variable y_α in the higher-order language.[10] The first-order approach, on the other hand, needs just one primitive predicate C^2 to represent the natural language predicate 'contemplate'. In fact, the first-order approach satisfies Davidson's learnability requirement on all counts.

Finally, it seems inevitable (especially in connection with stubborn desiderata such as 3, 8, 9, 23, and 24) that higher-order theories will be considerably more complex than their first-order counterparts. This unnecessary complexity is another count against the higher-order approach.

For all these reasons, it would seem that from the point of view of formal logico-linguistic theory, one is better off using the first-order approach.

Summing up, then, I have the following conclusion. The best representation of intensionality in natural language is provided by a first-order quantificational language that is fitted out with (defined or undefined) complex singular terms such as $[A]_{v_1 \ldots v_m}$, and depending on the value of m, these complex terms are semantically correlated with properties, relations, or propositions. A corollary of this conclusion is one of the underlying tenets of the book, the tenet that the theory of PRPs is part of the logic for natural language and as such is part of logic *per se* with all the attendant privileges and responsibilities.

With this matter at least provisionally settled I am finally ready for the first substantive task of this work, the formalization of intensional logic. The general strategy will be this. Given the above conclusions about the origin and character of intensionality, it follows that I shall have succeeded in formalizing intensional logic if I am successful in spelling out the logical properties of the special complex intensional terms $[A]_\alpha$. This is to be done in the two standard phases. First, I give a semantical characterization of the logical properties of the complex terms $[A]_\alpha$. Secondly, I attempt as nearly as possible to give a syntactical characterization of the same logical properties; this is done by the formulation of an axiomatic first-order intensional logic. Since these two tasks are independent of the question of whether the complex intensional terms $[A]_\alpha$ are defined or undefined, I am free to consider them as if they were undefined. By doing this, I am able to obtain one of the major results of the book, namely, that this first-order intensional logic is sound and complete and, hence, that the syntactic characterization of intensional logic is perfectly equivalent to the logically prior semantical characterization. It is after obtaining these results that I will look at the question of whether these complex intensional expressions can be defined and, in particular, whether they can be defined in a first-order extensional language.

Before moving on to the substantive tasks of the work, however, I want to make a brief digression on the topic of quantifying-in, which was considered in §7.

11. Quine and Church on Quantifying-in*

In §31 of *Word and Object* Quine proposes a way to represent quantifying-in that is rather different from the one I proposed in §7. At the heart of Quine's treatment is a multiplication of the senses of 'believe'. For example, Quine would provisionally represent the intuitively valid argument

x believes that x believes something.

∴ There is someone v such that x believes that v believes something.

in the following alternative manner:[11]

$$\frac{B^3(x, x, [(\exists u)B^2(x, u)]_x)}{\therefore\ (\exists v)B^3(x, v, [(\exists u)B^2(x, u)]_x).}$$

On analogy, then, Quine would represent the intuitively valid argument

x believes that x believes y.

∴ There is someone v and something u such that x believes that v believes u.

as follows:

$$\frac{B^4(x, x, y, [B^2(x, y)]_{xy})}{\therefore\ (\exists u, v)B^4(x, u, v, [B^2(x, y)]_{xy}).}$$

The important thing to notice is that three separate senses of 'believe'—represented by B^2, B^3, and B^4—have already been posited. Since for arbitrarily high numbers n, there are 'that'-clauses containing n distinct externally quantifiable variables, Quine's approach leads to infinitely many primitive 'belief'-predicates—B^2,

* The reader may skip over this section without losing the basic line of development of the book.

$B^3, B^4, \ldots, B^n, \ldots$—and, hence, to a violation of desideratum 13, Davidson's learnability requirement.

Now perhaps this problem can be remedied simply by getting rid of $B^4, B^5, B^6, \ldots$ and by doing their work with finite sequences plus the 'belief'-predicate B^3. There is, however, a further problem in the Quinean approach, a problem that has no easy remedy. Consider the following two formulas:

(1) For all y, if x believes y, then x believes that someone believes y.
(2) x believes y.

From (1) and (2) one can derive the following infinite list of formulas:

(3) x believes that someone believes y.
(4) x believes that someone believes that someone believes y.
(5) x believes that someone believes that someone believes that someone believes y.

....

In my bracket notation these derivations are represented simply as follows:

(1′) $(\forall y)(xB^2y \supset xB^2[(\exists u)uB^2y]^y)$
(2′) xB^2y
(3′) $\therefore\ xB^2[(\exists u)uB^2y]^y$ By (1′), (2′), UI, and MP
(4′) $\therefore\ xB^2[(\exists u)uB^2[(\exists u)uB^2y]^y]^y$ By (1′), (3′), UI, and MP
(5′) $\therefore\ xB^2[(\exists u)uB^2[(\exists u)uB^2[(\exists u)uB^2y]^y]^y]^y$ By (1′), (4′), UI, and MP

....

On the Quinean approach, by contrast, (1)–(3) would be represented as follows:

(1″) $(\forall y)(xB^2y \supset B^3(x, y, [(\exists u)uB^2y]_y)$
(2″) xB^2y
(3″) $\therefore\ B^3(x, y, [(\exists u)uB^2y]_y)$ By (1″), (2″), UI, and MP.

So far so good. However, it appears impossible on the present approach to go on to represent the derivation of (4) from (1) and (3). For there can be no instantiation of (1″) whose antecedent is (3″). The reason for this is that the antecedent of (1″) is an atomic sentence with two arguments whereas (3″) is an atomic sentence with three arguments. One might think that this problem can be circumvented by somehow using formulas containing B^3 in place of formulas containing B^2. But just try. All straightforward attempts to use this idea to expedite the above derivations just lead to further difficulties of their own. I leave it to the reader to convince himself of this.

My conclusion is that, if one adopts quantifier logic as his initial theoretical framework, there is no reasonable alternative to treating 'believes' as a univocal 2-place predicate and 'that'-clauses as defined or undefined singular terms in which externally quantifiable variables may occur. Indeed, if as I have recommended we use my bracket notation provisionally to represent 'that'-clauses, then the treatment that I believe Quine was looking for in *Word and Object* can, ironically, be achieved as follows:

$$[A]^{v_1 \ldots v_j} =_{\mathrm{df}} \langle\langle v_1, \ldots, v_j\rangle, [A]_{v_1 \ldots v_j}\rangle.$$

Although painfully unnatural, this treatment avoids all the syntactic difficulties that beset Quine's actual treatment.[12] For example, the derivation of (3), (4), ... from (1) and (2) can be represented as follows:

(1‴) $(\forall y)(xB^2y \supset xB^2\langle\langle y\rangle, [(\exists u)uB^2y]_y\rangle)$
(2‴) xB^2y

(3‴) $\therefore\ xB^2\langle\langle y\rangle, [(\exists u)uB^2y]_y\rangle$
By (1‴), (2‴), UI, and MP
(4‴) $\therefore\ xB^2\langle\langle\langle y\rangle, [(\exists u)uB^2y]_y\rangle, [(\exists u)uB^2y]_y\rangle$
By (1‴), (3‴), UI, and MP

....

The important thing to notice, however, is that this treatment, unlike Quine's official treatment, takes 'believes' to be a 2-place predicate and 'that'-clauses to be singular terms that may contain externally quantifiable variables. Thus, the conclusions reached in §§6–7 are sustained. This is all that I wanted to show here.

I will next make a few remarks about the inability of the Frege-

Church approach to adequately represent quantifying-in. This fact is not widely recognized and, therefore, deserves discussion. Consider the following formula:

(6) x believes that y is a spy.

Is it possible to represent this formula in Church's system as follows:

> There is an individual concept y_{ι_1} such that y_{ι_1} is a concept of the individual y_ι and x_ι believes the proposition that is the value of the *spy* sense-function when applied to the argument y_{ι_1}.

i.e.,

$(\exists y_{\iota_1})(y_\iota \,\Delta\, y_{\iota_1} \;\&\; B_{oo_1\iota}(x_\iota, S_{o_1\iota_1}(y_{\iota_1})))$?

The answer is negative. To see why, consider the following related sentence:

(7) Someone is the F and x believes that the F is a spy.

Intuitively, this sentence has two logically independent readings, an "opaque" reading and a "transparent" reading. In my bracket notation these two readings can be represented, respectively, by the following:

(8) $(\exists y)(y = (\iota z)(Fz) \;\&\; xB^2[S(\iota z)(Fz)])$
(9) $(\exists y)(y = (\iota z)(Fz) \;\&\; xB^2[Sy]^y)$

where the definite description $(\iota z)(Fz)$ has narrow scope. The opaque reading of (7) is represented in Church's system as follows:

(10) $(\exists y_\iota)(y_\iota = \iota_{\iota(o\iota)}(F_{o\iota}) \;\&\; B_{oo_1\iota}(x_\iota, S_{o_1\iota_1}(\iota_{\iota_1(o_1\iota_1)}(F_{o_1\iota_1}))))$.

This seems to be acceptable. However, suppose that the method suggested earlier for representing quantifying-in within Church's system were adopted. Then, it should be possible to represent the transparent reading of (7) with something like the following:

(11) $(\exists y_\iota)(\exists y_{\iota_1})(y_\iota = \iota_{\iota(o\iota)}(F_{o\iota}) \;\&\; y_{\iota_1} = \iota_{\iota_1(o_1\iota_1)}(F_{o_1\iota_1})$
$\;\&\; y_\iota \,\Delta\, y_{\iota_1} \;\&\; B_{oo_1\iota}(x_\iota, S_{o_1\iota_1}(y_{\iota_1})))$.

However, given the intended interpretation of Δ, $\iota_{\iota(o\iota)}$, $\iota_{\iota_1(o_1\iota_1)}$, $F_{o\iota}$, and $F_{o_1\iota_1}$, the following is a logical truth:

(12) $(\exists y_\iota)(y = \iota_{\iota(o\iota)}(F_{\iota(o\iota)})) \supset \iota_{\iota(o\iota)}(F_{o\iota}) \,\Delta\, \iota_{\iota_1(o_1\iota_1)}(F_{o_1\iota_1})$.

It follows from this that (10) and (11) are logically equivalent. But the two readings of (7) are logically independent. Therefore, (11) cannot be an adequate representation of the transparent reading of (7).

Some advocates of the Frege-Church approach to intensional language are not at all disturbed by this sort of outcome, for they are basically skeptical about the legitimacy of quantifying-in, at least as it arises in connection with the usual examples. However, there are examples of quantifying-in that should move even hard-line advocates of the Frege-Church approach, examples that ought to be representable by every treatment of intensional language. The existence of this sort of example has not, as far as I know, been discussed in the literature.

The following intuitively valid argument illustrates the sort of example I have in mind:

(13) For all y, if x believes y, then x believes that someone believes y.
(14) x believes that A.
(15) $\therefore$ x believes that someone believes that A.

Unlike some of the examples of quantifying-in, this example requires no special education of one's intuitions. Indeed, it is unlikely that there is a reading of this argument according to which it is not intuitively valid. Using my bracket notation, I can represent the argument simply as follows:

$$(13') \quad (\forall y)(xB^2y \supset xB^2[(\exists z)zB^2y]^y)$$
$$(14') \quad xB^2[A]$$
$$(15') \quad \therefore\ xB^2[(\exists z)zB^2[A]] \qquad \text{By UI and MP.}$$

In contrast to this approach, the Frege-Church approach (as it stands) does not appear to provide any adequate representation of this intuitively valid argument.

To see what the problem is, consider the following candidate representations within Church's system. First, one might attempt to represent the argument as follows:

$$(13'') \quad (\forall p_{o_1})(\forall p_{o_2})(p_{o_1} \,\Delta\, p_{o_2} \supset (B_{oo_1\iota}(x_\iota, p_{o_1}) \supset B_{oo_1\iota}(x_\iota, (\exists y_{\iota_1})(B_{o_1o_2\iota_1}(y_{\iota_1}, p_{o_2})))))$$
$$(14'') \quad B_{oo_1\iota}(x_\iota, A_{o_1})$$
$$(15'') \quad \therefore\ B_{oo_1\iota}(x_\iota, (\exists y_{\iota_1})(B_{o_1o_2\iota_1}(y_{\iota_1}, A_{o_2}))).$$

True, the inference from (13″) and (14″) to (15″) is valid—or at least it is when supplemented with $A_{o_1} \Delta A_{o_2}$ as an additional premise. However, this way of representing the above argument is not adequate; for (13″) is too strong. To see why, let p_{o_1} be some proposition believed by x_ι and let (13″) be true. Then, for every concept p_{o_2} of p_{o_1}, the following would have to be true:

$$B_{o o_1 \iota}(x_\iota, (\exists y_{\iota_1})(B_{o_1 o_2 \iota_1}(y_{\iota_1}, p_{o_2}))).$$

But this is implausible. To dramatize the implausibility, consider an example given by Church in a different connection (p. 22 n., 'A Formulation'). Let p_{o_1} be the proposition that it is necessary that everything has some property or other. This proposition is in fact the proposition mentioned on lines 27–8 of page 272 of Lewis and Langford's *Symbolic Logic*. Consider the following two sentences:

(16) x_ι believes that someone believes that it is necessary that everything has some property or other.

(17) x_ι believes that someone believes the proposition mentioned on lines 27–8 of page 272 of Lewis and Langford's *Symbolic Logic*.

Clearly, there is a reading of (16) and a reading of (17) according to which it is possible for (16) to be true when (17) is false. In my bracket notation these readings of (16) and (17) would be represented as follows:

(16′) $xB^2[(\exists y)yB^2[N^1[(\forall u)(\exists v)(u \Delta v)]]]$

(17′) $xB^2[(\exists y)yB^2(\iota w)(M^1 w)]$

where M^1 represents 'is mentioned on lines 27–8 of page 272 of Lewis and Langford's *Symbolic Logic*' and $(\iota w)(M^1 w)$ is a definite description having narrow scope. Let us keep these readings in mind. Now suppose that x believes our proposition p_{o_1}. In this event, (13) requires that (16) on the indicated reading is true; (13), however, does not require that (17) on the indicated reading is true. By contrast, (13″) requires that on the indicated readings both (16) and (17) are true. Thus, (13″) does not adequately represent (13); it is too strong.

The following is a second attempt to represent the inference from (13) and (14) to (15) within Church's system:

$$
\begin{array}{ll}
(13''') & (\forall p_{o_1})(B_{oo_1\iota}(x_\iota, p_{o_1}) \supset (\exists p_{o_2})(p_{o_1} \,\Delta\, p_{o_2} \\
 & \qquad\qquad \& \; B_{oo_1\iota}(x_\iota, (\exists y_{\iota_1})(B_{o_1o_2\iota_1}(y_{\iota_1}, p_{o_2}))))) \\
(14''') & \underline{B_{oo_1\iota}(x_\iota, A_{o_1})} \\
(15''') & \therefore\; (\exists p_{o_2})(A_{o_1} \,\Delta\, p_{o_2} \;\&\; B_{oo_1\iota}(x_\iota, (\exists y_{\iota_1})(B_{o_1o_2\iota_1}(y_{\iota_1}, p_{o_2})))).
\end{array}
$$

Although this argument is valid, it too fails to adequately represent the original. The reason for this is that the Churchian representation of (13) is now too weak. To see why, suppose that x_ι believes the proposition p_{o_1} discussed earlier, and suppose further that (16) is false on the reading isolated above. It follows that (13) is false as well. But notice that on the readings isolated above it is possible for (17) to be true even when (16) is false. In addition, if (17) on this reading is true, then so is (13'''). Therefore, from the fact that (13) is false it does not follow that (13''') is false. Hence, (13''') fails to adequately represent (13); it is too weak.

There is a third strategy by which one could attempt to represent the inference from (13) and (14) to (15) within Church's system. Namely, one could incorporate the modified Quinean treatment that I described earlier. According to this rather artificial treatment, objects of belief are identified with ordered pairs:

$$[A]^{v_1 \ldots v_j} =_{\mathrm{df}} \langle\langle v_1, \ldots, v_j\rangle, [A]_{v_1 \ldots v_j}\rangle.$$

However, incorporating this treatment within Church's system not only would violate the spirit of the Frege-Church theory but also would generate excessive complications in connection with the matter of type restrictions. It should be noted, moreover, that such a treatment would be inconsistent with the principle of identity underlying Church's Alternative (2), namely, the principle that necessary equivalence is sufficient for identity. To see this, note that the following is intuitively valid:

$$(\forall x)(\forall y)N^1[x = x \equiv y = y]^{xy}.$$

Therefore, given the principle of identity underlying Church's alternative 2, the following should also be valid:

$$(\forall x)(\forall y)[x = x]^x = [y = y]^y.$$

However, on the modified Quinean treatment, this sentence is definitionally equivalent to

$$(\forall x)(\forall y)\langle\langle x\rangle, [x = x]_x\rangle = \langle\langle y\rangle, [y = y]_y\rangle.$$

But the latter sentence is invalid, for if $x \neq y$, then

$$\langle\langle x\rangle, [x = x]_x\rangle \neq \langle\langle y\rangle, [y = y]_y\rangle.$$ [13]

From the foregoing criticisms it does not follow that there is no way to construct a unified representation of quantifying-in within Church's systems. However, no such unified representation suggests itself.[14]

Incidentally, before winding up these comments on quantifying-in, I should note that, if one were to attempt to develop a treatment of quantifying-in within Carnap's framework or Scheffler's framework, problems involving multiple embeddings of 'that'-clauses would arise. However, it appears that, by adapting the artificial modified Quinean treatment that I described above, one could surmount these problems at least formally. On the assumption that this is so, I have given Carnap and Scheffler '+' grades for desideratum 5 on the chart in §4.

I hope that this digression on alternate treatments of quantifying-in has helped to bring out the virtues of my bracket notation for representing quantifying-in. However, it is now time to leave these philosophical issues behind and to commence the study of formal intensional logic.

2

Intensional Logic

Intensional logic has never been completely and adequately formulated. To be convinced of this, consider two representative arguments:

Whatever x believes y believes.
x believes that A.
___
∴ y believes that A.

Being a bachelor is the same thing as being an unmarried man.
It is necessary that all and only bachelors are bachelors.
___
∴ It is necessary that all and only bachelors are unmarried men.

Neither of these intuitively valid arguments is even expressible in standard first-order predicate logic, even when epistemic and modal operators are adjoined. And while it is true that both of these arguments can be expressed in certain higher-order intensional logics, such higher-order logics are essentially incomplete, to mention just one of their shortcomings. But things are better than they might seem. When an intensional abstraction operation is adjoined to first-order logic, the result is an intensional logic that is equipped to represent the above arguments—and indeed, nearly all problematic intensional arguments. At the same time, unlike higher-order intensional logics, this first-order intensional logic is, surprising as it might seem, provably complete.

In what follows I will show how to construct such a logic. The construction requires the development of both a new formal language and a new semantic method. The new semantic method does not appeal to possible worlds, even as a heuristic. The heuristic used is simply that of properties, relations, and propositions, taken at face value. And unlike the various possible-worlds approaches to intensional logic, the approach developed here is adequate for treating both modal and intentional matters. Initially, the intensional logic will have two parts, one for each of the traditional conceptions of PRPs identified in §2. At the end of this chapter the two parts will be integrated.

12. A Formal Intensional Language

I begin by specifying the syntax for a first-order language with intensional abstraction. This language will be called L_ω. Primitive symbols:

Logical operators:	$\&, \neg, \exists$
Predicate letters:	$F_1^1, F_2^1, \ldots, F_p^q$
Variables:	$x, y, z, \ldots$
Punctuation:	(,), [,].

Simultaneous inductive definition of *term* and *formula* of L_ω:

(1) All variables are terms.
(2) If $t_1, \ldots, t_j$ are terms, then $F_i^j(t_1, \ldots, t_j)$ is a formula.
(3) If A and B are formulas and v_k a variable, then $(A \,\&\, B)$, $\neg A$, and $(\exists v_k)A$ are formulas.
(4) If A is a formula and $v_1, \ldots, v_m$, $0 \leq m$, distinct variables, then $[A]_{v_1 \ldots v_m}$ is a term.

In the limiting case where $m = 0$, $[A]$ is a term. All and only formulas and terms are *well-formed* expressions. An occurrence of a variable v_i in a well-formed expression is *bound* (*free*) if and only if it lies (does not lie) within a formula of the form $(\exists v_i)A$ or a term of the form $[A]_{v_1 \ldots v_i \ldots v_m}$. A variable is *free* (*bound*) in a well-formed expression if and only if it has (does not have) a free occurrence in that well-formed expression. A *sentence* is a formula having no free variables. The predicate letter F_1^2 is singled out as a distinguished logical predicate, and formulas of the form $F_1^2(t_1, t_2)$ are to be rewritten in the form $t_1 = t_2$. $\forall$, $\supset$, $\supset_{v_i \ldots v_j}$, $\vee$, $\equiv$, $\equiv_{v_i \ldots v_j}$ are to be defined in terms of $\exists$, &, and $\neg$ in the usual way. If v_i occurs free in A and is not one of the variables in the sequence of variables α, then v_i is an externally quantifiable variable in the term $[A]_\alpha$. Let the sequence δ be, in order, the externally quantifiable variables in $[A]_\alpha$; then $[A]_\alpha$ will sometimes be rewritten as $[A]_\alpha^\delta$ so that these variables can be identified at a glance.

Some observations are in order. First, on the intended informal interpretation of L_ω, a singular term $[A]_{v_1 \ldots v_m}$ denotes a proposition if $m = 0$, a property if $m = 1$, and an m-ary relation-in-intension if $m \geq 2$. Secondly, L_ω differs from a standard first-order language only in having these singular terms $[A]_{v_1 \ldots v_m}$. Thirdly, L_ω has a finite number of primitive constants, and hence, it satisfies desideratum 13, Davidson's learnability requirement. Of course, for purely mathematical purposes, one is free to adjoin an infinite number of additional primitive constants to L_ω. Yet if Davidson is right, such infinitistic extensions of L_ω will not qualify as idealized representations of natural language. Fourthly, L_ω contains no primitive names. My strategy with regard to names will be to proceed in two stages. First, I will study the logic of intensional language without names; that is, I will study the logic of L_ω as its stands. Once this task is completed, I will take up the question of how to treat names. There are two main competing theories of names—Frege's theory and Mill's theory. According to Frege's theory, names have descriptive content; according to Mill's theory, they do not. In §§38–9 it is shown that, given either theory, names can be successfully treated in the setting of L_ω. And finally, L_ω contains no functional constants: these are superfluous in L_ω since they can be contextually defined in terms of = and appropriate auxiliary predicates.[1]

Now let us reconsider the intuitively valid arguments mentioned at the outset of the chapter. In L_ω they can be represented as follows:

$$\begin{array}{l} (\forall z)(B(x, z) \supset B(y, z)) \\ B(x, [A]) \\ \hline \therefore\ B(y, [A]) \end{array}$$

$$\begin{array}{l} [B(x)]_x = [U(x) \,\&\, M(x)]_x \\ N([(\forall x)(B(x) \equiv B(x))]) \\ \hline \therefore\ N([(\forall x)(B(x) \equiv (U(x) \,\&\, M(x)))]).* \end{array}$$

Of course, to guarantee that these and other intuitively valid arguments come out valid in L_ω, I must first specify the semantics for L_ω.

* In order to enhance readability, I take the liberty here and elsewhere to use predicate letters (with or without indices) that do not strictly speaking belong to L_ω, and I occasionally delete some parentheses and commas.

13. A New Semantic Method

By what means should one characterize the semantics for L_ω? Since the aim is simply to characterize the logically valid formulas of L_ω, it will suffice to construct a Tarski-style definition of logical validity for L_ω. Such a definition will be built on Tarski-style definitions of truth for L_ω. The latter definitions will in turn depend in part on specifications of the denotations of the singular terms in L_ω. As already indicated, every formula of L_ω is just like a formula in a standard first-order extensional language except perhaps for the singular terms occurring in it. Therefore, once one has found a method for specifying the denotations of the singular terms of L_ω, the Tarski-style definitions of truth and validity for L_ω may be given in the customary way. What is being sought specifically is a method for characterizing the denotations of the singular terms of L_ω in such a way that a given singular term $[A]_{v_1 \ldots v_m}$ will denote an appropriate property, relation, or proposition, depending on the value of m.

Since L_ω has infinitely many complex singular terms $[A]_\alpha$, what is called for is a recursive specification of the denotation relation for L_ω. To do this I will arrange these singular terms into an order according to their syntactic kind and complexity. So, for example, just as the complex formula $((\exists x)Fx \,\&\, (\exists y)Gy)$ is the conjunction of the simpler formulas $(\exists x)Fx$ and $(\exists y)Gy$, I will say that the complex term $[(\exists x)Fx \,\&\, (\exists y)Gy]$ is the conjunction of the simpler terms $[(\exists y)Fx]$ and $[(\exists y)Gy]$. Similarly, just as the complex formula $\neg(\exists x)Fx$ is the negation of the simpler formula $(\exists x)Fx$, I will say that the complex term $[\neg(\exists x)Fx]$ is the negation of the simpler term $[(\exists x)Fx]$. The following are other examples: $[Rxy]_{yx}$ is the conversion of $[Rxy]_{xy}$; $[Sxyz]_{xzy}$ is the inversion of $[Sxyz]_{xyz}$; $[Rxx]_x$ is the reflexivization of $[Rxy]_{xy}$; $[Fx]_{xy}$ is the expansion of $[Fx]_x$; $[(\exists x)Fx]$ is the existential generalization of $[Fx]_x$; $[Fy]^y$ is the absolute predication of $[Fx]_x$ of y; $[F[Guvw]_{uvw}]$ is the absolute predication of $[Fx]_x$ of $[Guvw]_{uvw}$; $[F[Guvw]^w_{uv}]_w$ is the unary relativized predication of $[Fx]_x$ of $[Guvw]_{uvw}$; $[F[Guvw]^{vw}_u]_{vw}$ is the binary relativized predication of $[Fx]_x$ of $[Guvw]_{uvw}$; $[F[Guvw]^{uvw}]_{uvw}$ is the ternary relativized predication of $[Fx]_x$ of $[Guvw]_{uvw}$, and so on. In this way I isolate the following syntactic operations on intensional abstracts: conjunction, negation, conversion, inversion, reflexivization, expansion, existential generalization, absolute predication, unary relativized predication, binary relativized predication, ..., n-ary relativized predication, ...[2]

Those intensional abstracts whose form is $[F_h^m(v_1, \ldots, v_m)]_{v_1 \ldots v_m}$ are syntactically simpler than all others. I will call them *elementary*. And the denotation of an elementary intensional abstract $[F_h^m(v_1, \ldots, v_m)]_{v_1 \ldots v_m}$ is just the property or relation expressed by the primitive predicate F_h^m. The denotation of a more complex abstract $[A]_\alpha$ is defined in terms of the denotation(s) of the relevant syntactically simpler abstract(s). However, to state this definition, one must have a general technique for modeling PRPs.

Suppose that one were to use one of the previous approaches to this subject—namely, the approach of Russell, of Church, or of the possible-worlds theorists Montague, Kaplan, D. Lewis, *et al.* In that case one would be led to identify properties and relations with certain functions. I find such identification unintuitive. (The taste of pineapple, the missing shade of blue—are these functions?) Furthermore, the identification of properties and relations with functions leads naturally—and perhaps inevitably—to a hierarchy of artificially restricted logical types. (See desideratum 14, §4.) Since the thesis that properties and relations are functions is linked in this way to type theory, it proves to be more compatible with the higher-order approach to the logic of PRPs than it is with the first-order approach. In a first-order setting, such as that provided by L_ω, the identification of properties and relations with functions generates unwanted and unnecessary complications and restrictions. The alternative is to take properties and relations, as well as propositions, at face value, i.e., as real, irreducible entities. This is what I will do.

The identification of intensional entities with functions lies at the heart of the possible-worlds semantic method. If, as I have proposed, intensional entities are taken at face value and not as covert functions, then the possible-worlds semantic method will be of no use to us. But how, then, is the denotation of a given complex term $[A]_\alpha$ to be determined from the denotation(s) of the relevant syntactically simpler term(s)? My answer is that the new denotation is determined *algebraically*. That is, the new denotation is determined by the application of the relevant *fundamental logical operation* to the denotation(s) of the relevant syntactically simpler term(s). Let me explain.

Consider the following propositions, for example: $[(\exists x)Fx]$, $[(\exists y)Gy]$, $[(\exists x)Fx \mathbin{\&} (\exists y)Gy]$. (Note: in this paragraph and the next I will be using—not mentioning—terms from L_ω.) What is the most obvious logical relation holding among these propositions? Answer: the third proposition is the conjunction of the first two. Similarly, what is the most obvious logical relation among the properties $[Fx]_x$, $[Gx]_x$, and $[Fx \mathbin{\&} Gx]_x$? As before, the third is the conjunction of the first two. And what is the most obvious logical relation holding between the propositions $[(\exists x)Fx]$ and $[\neg(\exists x)Fx]$? Answer: the second is the negation of the first. Similarly, what is the most obvious logical relation holding between the properties $[Fx]_x$ and $[\neg Fx]_x$? As before, the second is the negation of the first. In a like manner I arrive at the following fundamental logical relationships: $[Rxy]_{yx}$ is the converse of $[Rxy]_{xy}$; $[Sxyz]_{xzy}$ is the inverse of $[Sxyz]_{xyz}$; $[Rxx]_x$ is the reflexivization of $[Rxy]_{xy}$; $[Fx]_{xy}$ is the expansion of $[Fx]_x$; $[(\exists x)Fx]$ is the existential generalization of $[Fx]_x$; $[Fy]^y$ is the absolute predication of $[Fx]_x$ of y; $[F[Guvw]_{uvw}]$ is the absolute predication of $[Fx]_x$ of $[Guvw]_{uvw}$; $[F[Guvw]^{w}_{uv}]_w$ is the unary relativized predication of $[Fx]_x$ of $[Guvw]_{uvw}$; $[F[Guvw]^{vw}_{u}]_{vw}$ is the binary relativized predication of $[Fx]_x$ of $[Guvw]_{uvw}$; $[F[Guvw]^{uvw}]_{uvw}$ is the ternary relativized predication of $[Fx]_x$ of $[Guvw]_{uvw}$, and so on. Thus, in one-to-one correspondence with the earlier syntactic operations on intensional abstracts there are fundamental logical operations on intensional entities: conjunction, negation, conversion, ...

The first two fundamental logical operations are intensional analogues of the two operations from Boolean algebra. A Boolean algebra having two elements (T and F) is an extensional model of first-order sentential logic. The next four operations are intensional analogues of operations from the algebra of relations, whose origins are found in the work of Peirce and Schröder. The algebra of relations, or transformation algebra as it is called, is the algebra for extensional relations. A transformation algebra is an extensional model of first-order predicate logic without quantifiers. The next operation, existential generalization, is an intensional analogue of the special new operation found in polyadic algebra. Polyadic algebra is just the algebra for extensional relations *with* quantification. A polyadic algebra is an extensional model of first-order predicate logic with quantifiers.[3] Finally, the predication operations, absolute predication and n-ary relativized predication, $n \geq 1$, are further operations that I have isolated for the purpose of modeling first-order quantifier logic with distinguished singular terms, including in particular intensional abstracts. Absolute predication is straightforward. As indicated above, the absolute predication of $[Fx]_x$ of y is $[Fy]^y$, i.e., the proposition that y is F. Similarly, the absolute predication of $[Fx]_x$ of $[Gy]_y$ is $[F[Gy]_y]$, i.e., the proposition that the property of being G is F. Relativized predication differs somewhat from absolute predication. It also predicates a property of an intension, but it involves in addition a simultaneous predication of which that intension is the result. So, for example, the unary relativized predication of $[Fx]_x$ of $[Gy]_y$ is $[F[Gy]^y]_y$, i.e., the property of being something y such that the proposition that y is G is F. To give a concrete example, the unary relativized predication of the property *being believed* of the property *being a spy* is the property *being believed to be a spy*. The other relativized predication operations behave analogously; of course, their second arguments must be intensions of appropriately higher degree.

Taken together, these fundamental logical operations have the following property. Choose any intensional abstract $[A]_\alpha$ in L_ω that is not elementary. If $[A]_\alpha$ is obtained from $[B]_\beta$ via the syntactic operation of negation (conversion, inversion, reflexivization, expansion, existential generalization), then the denotation of $[A]_\alpha$ is the result of applying the logical operation of negation (conversion, inversion, reflexivization, expansion, existential generalization) to the denotation of $[B]_\beta$. The same thing holds *mutatis mutandis* for abstracts that, syntactically, are conjunctions or predications (absolute or relativized). In this way, therefore, these fundamental logical operations make it possible to define recursively the denotation relation for all of the complex intensional abstracts $[A]_\alpha$ in L_ω.

The algebraic semantics for L_ω is thus to be specified in stages. First, an algebra of properties, relations, and propositions—or an algebraic model structure, as I will call it—is posited. Secondly, an intensional interpretation of the primitive predicates is given. Thirdly, the denotation relation for the terms of L_ω is recursively defined. Fourthly, the notion of truth for formulas is defined. Finally, in the customary Tarski fashion, the notion of logical validity for formulas of L_ω is defined.

Now a structure β is a Boolean algebra if and only if (i) β is an ordered set consisting of a universe or domain $\mathscr{D}$ and two operations on $\mathscr{D} \times \mathscr{D}$ and $\mathscr{D}$, respectively, and (ii) the elements of β satisfy certain specifiable conditions. By analogy, $\mathscr{M}$ is an algebraic model structure if and only if (i) $\mathscr{M}$ is an ordered set consisting of a universe or domain $\mathscr{D}$ and the fundamental logical operations on $\mathscr{D} \times \mathscr{D}, \mathscr{D}, \ldots$, respectively (plus certain supplementary elements), and (ii) the elements of $\mathscr{M}$ satisfy certain specifiable conditions. In §2 I mentioned that historically there have been two competing conceptions of intensional entities. According to conception 1, intensional entities are identical if and only if they are necessarily equivalent. According to conception 2, each definable intensional entity is such that, when it is defined completely, it has a unique, non-circular definition. By suitably adjusting the conditions imposed on the elements of a given algebraic model structure $\mathscr{M}$, one can fix the exact character of the intensional entities that $\mathscr{M}$ is designed to model. In particular, by suitably formulating the conditions imposed on the elements of $\mathscr{M}$, one can make precise what it takes for the intensional entities modeled by $\mathscr{M}$ to conform to conception 1 or conception 2.

In this way one actually arrives at two distinct types of algebraic model structures—type 1 and type 2. In turn, one arrives at two distinct notions of logical validity for L_ω—validity$_1$ and validity$_2$, i.e., truth-in-all-type-1-model-structures and truth-in-all-type-2-model-structures.

With these preliminary remarks in mind I will now use the new semantic method to lay out in detail the formal semantics for L_ω.

14. The Formal Semantics*

Algebraic model structures. An algebraic model structure (or model structure, for short) is any structure

$$\langle \mathscr{D}, \mathscr{P}, \mathscr{K}, \mathscr{G}, \mathrm{Id}, \mathrm{Conj}, \mathrm{Neg}, \mathrm{Exist}, \mathrm{Exp}, \mathrm{Inv}, \\ \mathrm{Conv}, \mathrm{Ref}, \mathrm{Pred}_0, \mathrm{Pred}_1, \ldots, \mathrm{Pred}_k, \ldots \rangle$$

whose elements simultaneously satisfy the conditions set forth below. $\mathscr{D}$ is the domain of discourse and is non-empty. $\mathscr{P}$ is a relation on $\mathscr{D}$ that serves to partition $\mathscr{D}$ into a denumerable number of disjoint subdomains: $\mathscr{D}_{-1}, \mathscr{D}_0, \mathscr{D}_1, \mathscr{D}_2, \mathscr{D}_3, \ldots$.[4] The elements

* Readers seeking a quick overview may skip this section.

of $\mathscr{D}_{-1}$ are to be thought of as particulars; the elements of $\mathscr{D}_0$, as propositions; the elements of $\mathscr{D}_1$, as properties, and the elements of $\mathscr{D}_i$, for $i \geq 2$, as i-ary relations. Although $\mathscr{D}_i$, $i \geq 0$, may not be empty, I do permit $\mathscr{D}_{-1}$ to be empty. $\mathscr{K}$ is a set of functions on $\mathscr{D}$. These functions are to be thought of as telling us the alternate or possible extensions of the elements of $\mathscr{D}$. Specifically, they tell us that the extension of a particular is itself, that the extension of a proposition is a truth value, that the extension of a property is a subset of $\mathscr{D}$, and that the extension of an i-ary relation is a set of ordered i-tuples of members of $\mathscr{D}$.[5] Thus, for $H \in \mathscr{K}$ and $x \in \mathscr{D}$, the following hold: if $x \in \mathscr{D}_{-1}$, then $H(x) = x$; if $x \in \mathscr{D}_0$, then $H(x) = \mathrm{T}$ or $H(x) = \mathrm{F}$; if $x \in \mathscr{D}_1$, then $H(x) \subseteq \mathscr{D}$; if, for $i > 1$, $x \in \mathscr{D}_i$, then $H(x) \subseteq {}^i\mathscr{D}$. The next element of a model structure is the function $\mathscr{G}$. $\mathscr{G}$ is a distinguished element of $\mathscr{K}$ and is to be thought of as the function that determines the *actual* extensions of the elements of $\mathscr{D}$. The element Id of a model structure is a distinguished element of $\mathscr{D}_2$ and is thought of as the fundamental logical relation-in-intension *identity*. Id must satisfy the following condition:

$$(\forall H \in \mathscr{K})(H(\mathrm{Id}) = \{xy \in \mathscr{D} : x = y\}).$$

That is, for every $H \in \mathscr{K}$, H singles out the extensional identity relation on $\mathscr{D}$ to be the extension of the intensional identity relation Id. The remaining elements of a model structure are functions which are thought of as fundamental logical operations on intensional entities. The domains and ranges of these operations are as follows:[6]

1. Conj: $\mathscr{D}_i \times \mathscr{D}_i \underset{\text{into}}{\rightarrow} \mathscr{D}_i$ for each $i \geq 0$

2. Neg: $\mathscr{D}_i \underset{\text{into}}{\rightarrow} \mathscr{D}_i$ for each $i \geq 0$

3. Exist: $\mathscr{D}_i \underset{\text{into}}{\rightarrow} \mathscr{D}_{i-1}$ for $i \geq 1$

   $\mathscr{D}_0 \underset{\text{into}}{\rightarrow} \mathscr{D}_0$

4. Exp: $\mathscr{D}_i \underset{\text{into}}{\rightarrow} \mathscr{D}_{i+1}$ for $i \geq 0$

5. Inv: $\mathscr{D}_i \underset{\text{into}}{\rightarrow} \mathscr{D}_i$ for $i \geq 3$

6. Conv: $\mathscr{D}_i \underset{\text{into}}{\rightarrow} \mathscr{D}_i$ for $i \geq 2$

7. Ref: $\mathscr{D}_i \underset{\text{into}}{\rightarrow} \mathscr{D}_{i-1}$ for $i \geq 2$

8.0 Pred_0: $\mathscr{D}_i \times \mathscr{D} \underset{\text{into}}{\rightarrow} \mathscr{D}_{i-1}$ for $i \geq 1$

8.1 Pred_1: $\mathscr{D}_i \times \mathscr{D}_j \underset{\text{into}}{\rightarrow} \mathscr{D}_i$ for $i, j \geq 1$

8.2 Pred_2: $\mathscr{D}_i \times \mathscr{D}_j \underset{\text{into}}{\rightarrow} \mathscr{D}_{i+1}$ for $i \geq 1$ and $j \geq 2$

8.3 Pred_3: $\mathscr{D}_i \times \mathscr{D}_j \underset{\text{into}}{\rightarrow} \mathscr{D}_{i+2}$ for $i \geq 1$ and $j \geq 3$

....[7]

The following conditions specify how the extensions of elements in $\mathscr{D}$ are affected by each of these operations. For all $H \in \mathscr{K}$ and all $u, v, x_1, \ldots, x_i, x_{i+1}, y_1, \ldots, y_k \in \mathscr{D}$:

1. $H(\text{Conj}(u, v)) = \text{T} \equiv$
   $(H(u) = \text{T} \;\&\; H(v) = \text{T})$ (for $u, v \in \mathscr{D}_0$)

   $\langle x_1, \ldots, x_i \rangle \in H(\text{Conj}(u, v)) \equiv$
   $(\langle x_1, \ldots, x_i \rangle \in H(u) \;\&\; \langle x_1, \ldots, x_i \rangle \in H(v))$
   (for $u, v \in \mathscr{D}_i, i \geq 1$)

2. $H(\text{Neg}(u)) = \text{T} \equiv H(u) = \text{F}$ (for $u \in \mathscr{D}_0$)

   $\langle x_1, \ldots, x_i \rangle \in H(\text{Neg}(u)) \equiv$
   $\langle x_1, \ldots, x_i \rangle \notin H(u)$ (for $u \in \mathscr{D}_i, i \geq 1$)

3. $H(\text{Exist}(u)) = \text{T} \equiv H(u) = \text{T}$ (for $u \in \mathscr{D}_0$)

   $H(\text{Exist}(u)) = \text{T} \equiv (\exists x_1)(x_1 \in H(u))$ (for $u \in \mathscr{D}_1$)

   $\langle x_1, \ldots, x_{i-1} \rangle \in H(\text{Exist}(u)) \equiv$
   $(\exists x_i)(\langle x_1, \ldots, x_{i-1}, x_i \rangle \in H(u))$ (for $u \in \mathscr{D}_i, i \geq 2$)

4. $x_1 \in H(\mathrm{Exp}(u)) \equiv H(u) = \mathrm{T}$ (for $u \in \mathscr{D}_0$)

$\langle x_1, \ldots, x_i, x_{i+1}\rangle \in H(\mathrm{Exp}(u)) \equiv$
$\langle x_1, \ldots, x_i\rangle \in H(u)$ (for $u \in \mathscr{D}_i, i \geq 1$)

5. $\langle x_1, \ldots, x_{i-2}, x_i, x_{i-1}\rangle \in H(\mathrm{Inv}(u)) \equiv$
$\langle x_1, \ldots, x_{i-2}, x_{i-1}, x_i\rangle \in H(u)$ (for $u \in \mathscr{D}_i, i \geq 3$)

6. $\langle x_i, x_1, \ldots, x_{i-1}\rangle \in H(\mathrm{Conv}(u)) \equiv$
$\langle x_1, \ldots, x_{i-1}, x_i\rangle \in H(u)$ (for $u \in \mathscr{D}_i, i \geq 2$)

7. $\langle x_1, \ldots, x_{i-1}\rangle \in H(\mathrm{Ref}(u)) \equiv$
$\langle x_1, \ldots, x_{i-1}, x_{i-1}\rangle \in H(u)$ (for $u \in \mathscr{D}_i, i \geq 2$)

8.0 $H(\mathrm{Pred}_0(u, y_1)) = \mathrm{T} \equiv y_1 \in H(u)$ (for $u \in \mathscr{D}_1$)

$\langle x_1, \ldots, x_{i-1}\rangle \in H(\mathrm{Pred}_0(u, y_1)) \equiv$
$\langle x_1, \ldots, x_{i-1}, y_1\rangle \in H(u)$ (for $u \in \mathscr{D}_i, i \geq 2$)

8.1 $\langle x_1, \ldots, x_{i-1}, y_1\rangle \in H(\mathrm{Pred}_1(u, v)) \equiv$
$\langle x_1, \ldots, x_{i-1}, \mathrm{Pred}_0(v, y_1)\rangle \in H(u)$
(for $u \in \mathscr{D}_i$, $i \geq 1$,
and $v \in \mathscr{D}_j$, $j \geq 1$)

8.2 $\langle x_1, \ldots, x_{i-1}, y_1, y_2\rangle \in H(\mathrm{Pred}_2(u, v)) \equiv$
$\langle x_1, \ldots, x_{i-1}, \mathrm{Pred}_0(\mathrm{Pred}_0(v, y_2), y_1)\rangle \in H(u)$
(for $u \in \mathscr{D}_i, i \geq 1$,
and $v \in \mathscr{D}_j, j \geq 2$)

....[8]

This completes the characterization of what a model structure is.

Type 1 Model Structures

A model structure is type 1 *iff*$_{\mathrm{df}}$ it satisfies the following auxiliary condition:

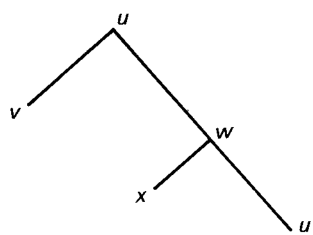
$$(\forall x, y \in \mathscr{D}_i)((\forall H \in \mathscr{K})(H(x) = H(y)) \supset x = y), \quad \text{for all } i \geq -1.$$

This condition provides us with a precise statement of conception 1.

Specifically, this condition rules out the possibility of there being two (or more) elements of any given subdomain $\mathscr{D}_i$ that are necessarily equivalent.

Type 2 Model Structures

A model structure is type 2 *iff*$_{\mathrm{df}}$ its operations Conj, Neg, Exist, Exp, Inv, Conv, Ref, Pred$_0$, Pred$_1$, Pred$_2$, ... are (i) one-one, (ii) disjoint in their ranges, and (iii) non-cycling. Auxiliary conditions (i)–(iii) provide us with a precise formulation of conception 2. For, taken together, (i) and (ii) guarantee that the action of the inverses of the fundamental logical operations in a given type 2 model structure $\mathscr{M}$ is to decompose the elements of $\mathscr{D}$ into unique (possibly infinite) trees. And condition (iii) insures that, for each item u in such a decomposition tree, u cannot occur on any path descending from u. So the following is the sort of situation ruled out by condition (iii):



Hence, whereas conditions (i) and (ii) insure that the elements of $\mathscr{D}$ have at most one complete definition in terms of the elements of $\mathscr{D}$ plus the fundamental logical operations, condition (iii) insures that such definitions are never circular.

Notice by the way that in the formal characterizations of what it is to be a type 1 or type 2 algebraic model structure no use is made of any of the following intuitive notions: particular, property, relation, proposition, alternative or possible extension, actual extension, complete definition. For what it is worth, type 1 and type 2 model structures are characterized formally in exclusively set-theoretic terms.[9]

At the close of §2, I mentioned that there are various intermediate conceptions of PRPs between conceptions 1 and 2. To model such intermediate conceptions, one need only appropriately adjust the auxiliary conditions imposed on algebraic model structures.

Consider, for example, the conception that is like conception 2 except that it imposes less strenuous identity conditions on conjunctions so that $[A\alpha \,\&\, B\alpha]_\alpha = [B\alpha \,\&\, A\alpha]_\alpha$ and $[(A\alpha \,\&\, B\alpha) \,\&\, C\alpha]_\alpha = [A\alpha \,\&\, (B\alpha \,\&\, C\alpha)]_\alpha$. The model structures appropriate to this conception are just like type 2 model structures except for the auxiliary conditions imposed on the conjunction operation Conj. Specifically, we exempt Conj from condition (i) and instead require that items in its range (i.e., conjunctions) can be decomposed under its inverse into a unique set of items (i.e., conjuncts) but in no special order. Accordingly, the inverse of Conj behaves rather like the operation of prime factorization in number theory: every natural number is factorable into a unique set of primes, yet there is no special order in which these prime factors must be multiplied in order to obtain the original number.

The field here is very rich. But conceptions 1 and 2 are the motherlode, and we should be happy to explore there for quite a while.

Truth and Validity

An interpretation $\mathscr{I}$ for L_ω relative to model structure $\mathscr{M}$ is a function that assigns to the predicate letter F_1^2 (i.e., $=$) the element $\mathrm{Id} \in \mathscr{M}$ and, for each predicate letter F_i^j in L_ω, assigns to F_i^j some element of the subdomain $\mathscr{D}_j \subset \mathscr{D} \in \mathscr{M}$. An assignment $\mathscr{A}$ for L_ω relative to model structure $\mathscr{M}$ is a function that maps the variables of L_ω into the domain $\mathscr{D} \in \mathscr{M}$. Truth $\mathrm{T}_{\mathscr{I}\mathscr{A}\mathscr{M}}$ is defined in terms of denotation $\mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}$, which will be defined subsequently:

$$\mathrm{T}_{\mathscr{I}\mathscr{A}\mathscr{M}}(A) \textit{ iff}_{\mathrm{df}}\ \mathscr{G}(\mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}([A])) = \mathrm{T}.$$

That is, formula A is true on interpretation $\mathscr{I}$ and assignment $\mathscr{A}$ relative to model structure $\mathscr{M}$ if and only if the actual extension of the proposition denoted by the term $[A]$ is the truth value T. (Of course, $\mathrm{T}_{\mathscr{I}\mathscr{A}\mathscr{M}}$ could instead be given a standard Tarski-style recursive definition (see lemma 6 in §15), but in the algebraic setting a direct definition suffices. A recursive definition is needed only in the definition of $\mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}$.) Then I define the two notions of validity for L_ω:

> a formula A is *valid*$_1$ *iff*$_{\mathrm{df}}$ for every type 1 model structure $\mathscr{M}$ and for every interpretation $\mathscr{I}$ and every assignment $\mathscr{A}$ relative to $\mathscr{M}$, $\mathrm{T}_{\mathscr{I}\mathscr{A}\mathscr{M}}(A)$.

a formula A is *valid*$_2$ *iff*$_{df}$ for every type 2 model structure $\mathcal{M}$ and for every interpretation $\mathcal{I}$ and every assignment $\mathcal{A}$ relative to $\mathcal{M}$, $T_{\mathcal{I}\mathcal{A}\mathcal{M}}(A)$.

Denotation

It remains to define the denotation function $D_{\mathcal{I}\mathcal{A}\mathcal{M}}$, which was referred to in the truth definition. To do this, I must first define the basic syntactic operations on intensional abstracts that were mentioned informally in §13.* I begin by introducing some preliminary syntactic notions.

I will say that a term $[A]_\alpha$ is *normalized* if and only if all the variables in the sequence of variables α occur free in A and α displays the order in which these variables first occur free in A. If a variable occurs free in more than one of the terms $t_1, \ldots, t_j$ in the atomic formula $F^j_i(t_1, \ldots, t_j)$, then this variable will be called a *reflected variable* in $F^j_i(t_1, \ldots, t_j)$. If the formula A is atomic and if the variables in the sequence of variables α are all free in A, then the term $[A]_\alpha$ will be called a *prime term*. If α contains a variable that is reflected in atomic formula A, then a prime term $[A]_\alpha$ will be called a *prime reflection term*. Let $[F^j_i(t_1, \ldots, [B]^\delta_\gamma, \ldots, t_j)]_\alpha$ be a prime term that is not a prime reflection term. Then, if some variable occurs in both α and δ, the prime term will be called a *prime relativized predication term*, and the variable will be called a *relativized variable*.

Every term $[A]_\alpha$ has associated with it a certain permutation of the variables in α that I will designate as *primary relative* to $[A]_\alpha$. (I admit the possibility that α itself can be primary relative to $[A]_\alpha$.) There are three cases.

Case (1): prime reflection terms $[A]_\alpha$. Suppose that α is some permutation of the sequence of variables $v_1, \ldots, v_p$ and that $[A]_{v_1 \ldots v_p}$ is normalized. Suppose further that, among the variables in α that are reflected in A, v_k is the one that has the right most free occurrence in A. In this case, the sequence $v_1, \ldots, v_{k-1}, v_{k+1}, \ldots, v_p, v_k$ is *primary relative* to $[A]_\alpha$.

* In doing this, I encounter certain intricacies, which arise because of the need to keep track of the various permutations of the subscripted variables α in the terms $[A]_\alpha$. Most of the intricacies could be avoided here by adopting the alternate technique developed in my 'Completeness in the Theory of Properties, Relations, and Propositions'. In any event my general algebraic approach is wedded to no particular treatment of this matter.

Case (*2*): $[A]_\alpha$ is a prime relativized predication term $[F_i^j(t_1, \ldots, [B]_\gamma^\delta, \ldots, t_j)]_\alpha$. Let α be a permutation of the sequence of variables $u_1, \ldots, u_p$, $v_1, \ldots, v_q$, $w_1, \ldots, w_r$ such that the latter sequence displays the order in which these variables first occur free in A. Let $[B]_\gamma^\delta$ be the left most argument of F_i^j containing relativized variables. Finally, let $v_1, \ldots, v_q$ be all such relativized variables in $[B]_\gamma^\delta$. Then the sequence $u_1, \ldots, u_p, w_1, \ldots, w_r, v_1, \ldots, v_q$ is *primary relative* to $[A]_\alpha$.

Case (*3*): $[A]_\alpha$ is neither a prime reflection term nor a prime relativized predication term. Let α be a permutation of the sequence of variables $v_1, \ldots, v_p, v_{p+1}, \ldots, v_{p+k}$ where $[A]_{v_1 \ldots v_p}$ is normalized and $v_{p+1}, \ldots, v_{p+k}$ are in order of their occurrence in α the variables not occurring free in A. Then the sequence $v_1, \ldots, v_p, v_{p+1}, \ldots, v_{p+k}$ is *primary relative* to $[A]_\alpha$. (I allow that $v_1, \ldots, v_p$ or $v_{p+1}, \ldots, v_{p+k}$ is an empty sequence.)

I am now prepared to define the basic syntactic operations on intensional abstracts of L_ω.

(1) If $[(A \,\&\, B)]_\alpha$ is normalized, it is the *conjunction* of $[A]_\alpha$ and $[B]_\alpha$.
(2) If $[\neg A]_\alpha$ is normalized, it is the *negation* of $[A]_\alpha$.
(3) Let $[(\exists v_k)A]_\alpha$ be normalized. Then, if v_k is free in A, $[(\exists v_k)A]_\alpha$ is the *existential generalization* of $[A]_{\alpha v_k}$; otherwise, $[(\exists v_k)A]_\alpha$ is the *existential generalization* of $[A]_\alpha$.
(4) If $[A]_\alpha$ is normalized and if v_{s+1} is the alphabetically earliest variable not occurring in $[A]_{\alpha v_1 \ldots v_s}$, then $[A]_{\alpha v_1 \ldots v_s v_{s+1}}$ is the *expansion* of $[A]_{\alpha v_1 \ldots v_s}$.
(5)–(6) Suppose that the sequence $v_1, v_2, \ldots, v_{s-1}, v_s$ is not primary relative to $[A]_{v_1 v_2 \ldots v_{s-1} v_s}$. Suppose instead that the sequence $u_1, \ldots, u_{s-1}, u_s$ is primary relative to $[A]_{v_1 v_2 \ldots v_{s-1} v_s}$. In this case if, for some $h, k \geq 1$, $u_1, \ldots, u_k = v_h, \ldots, v_{h+k-1}$ and $u_{k+1} = v_s \neq v_{h+k}$, then $[A]_{v_1 v_2 \ldots v_{s-1} v_s}$ is the *inversion* of $[A]_{v_1 v_2 \ldots v_s v_{s-1}}$; otherwise, $[A]_{v_1 v_2 \ldots v_{s-1} v_s}$ is the *conversion* of $[A]_{v_2 \ldots v_{s-1} v_s v_1}$.
(7) Let $[F_i^j(t_1, \ldots, t_k(v_r), \ldots, t_j)]_{\alpha v_r}$ be a prime reflection term relative to which the sequence αv_r is primary. Suppose that $t_k(v_r)$ is the right most argument of F_i^j in which the reflected variable v_r has a free occurrence. And suppose, finally, that v_s is the alphabetically earliest variable not occurring in

$t_1, \ldots, t_j$. Then $[F_i^j(t_1, \ldots, t_k(v_r), \ldots, t_j)]_{\alpha v_r}$ is the *reflexivization* of $[F_i^j(t_1, \ldots, t_k(v_s), \ldots, t_j)]_{\alpha v_r v_s}$.

(8) Suppose that $[F_i^j(t_1, \ldots, t_{k-1}, t_k, t_{k+1}, \ldots, t_j)]_\alpha$ is a normalized non-prime-reflection term. Let the terms $t_1, \ldots, t_{k-1}$ be variables all of which occur in the sequence α. Suppose that no variable that occurs free in t_k also occurs in the sequence α, and let v_r be the alphabetically earliest variable not occurring in $t_1, \ldots, t_j$. Then, $[F_i^j(t_1, \ldots, t_{k-1}, t_k, t_{k+1}, \ldots, t_j)]_\alpha$ is the *predication*$_0$ of $[F_i^j(t_1, \ldots, t_{k-1}, v_r, t_{k+1}, \ldots, t_j)]_{\alpha v_r}$ of t_k. Alternatively, let $[F_i^j(t_1, \ldots, t_{k-1}, [B]_\gamma^\delta, t_{k+1}, \ldots, t_j)]_{\alpha v_1 \ldots v_m}$ be a prime relativized predication term, where $m \geq 1$. Suppose that the terms $t_1, \ldots, t_{k-1}$ are variables that occur in the sequence α. And suppose that the sequence $\alpha, v_1, \ldots, v_m$ is primary relative to this prime relativized predication term and that $v_1, \ldots, v_m$ are the relativized variables occurring in $[B]_\gamma^\delta$. Then, this term is the *predication*$_m$ of $[F_i^j(t_1, \ldots, t_{k-1}, v_1, t_{k+1}, \ldots, t_j)]_{\alpha v_1}$ of $[B]_{\gamma v_1 \ldots v_m}^{\delta'}$, where δ' is the result of deleting the relativized variables $v_1, \ldots, v_m$ from δ.

For each non-elementary intensional abstract in L_ω, either it or one of its alphabetic variants[10] falls into the range of one of these syntactic operations, and no two non-elementary abstracts that are alphabetic variants fall into the range of more than one of them. In this sense, these operations serve to partition the class of non-elementary abstracts into denumerably many disjoint *syntactic kinds*: conjunctions, negations, existential generalizations, expansions, inversions, conversions, reflexivizations, predications$_0$, predications$_1$, predications$_2$, ... Using these notions, I inductively define the *denotation* function $D_{\mathscr{I}\mathscr{A}\mathscr{M}}$:

Variables: $D_{\mathscr{I}\mathscr{A}\mathscr{M}}(v_i) = \mathscr{A}(v_i)$

Elementary complex terms: $D_{\mathscr{I}\mathscr{A}\mathscr{M}}([F_i^j(v_1, \ldots, v_j)]_{v_1 \ldots v_j}) = \mathscr{I}(F_i^j)$

Non-elementary complex terms:

1. If t is the conjunction of r and s, then $D_{\mathscr{I}\mathscr{A}\mathscr{M}}(t) = \mathrm{Conj}(D_{\mathscr{I}\mathscr{A}\mathscr{M}}(r), D_{\mathscr{I}\mathscr{A}\mathscr{M}}(s))$.
2. If t is the negation of r, then $D_{\mathscr{I}\mathscr{A}\mathscr{M}}(t) = \mathrm{Neg}(D_{\mathscr{I}\mathscr{A}\mathscr{M}}(r))$.
3. If t is the existential generalization of r, then $D_{\mathscr{I}\mathscr{A}\mathscr{M}}(t) = \mathrm{Exist}(D_{\mathscr{I}\mathscr{A}\mathscr{M}}(r))$.

4. If t is an alphabetic variant of the expansion of r, then $\mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}(t) = \mathrm{Exp}(\mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}(r))$.
5. If t is the inversion of r, then $\mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}(t) = \mathrm{Inv}(\mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}(r))$.
6. If t is the conversion of r, then $\mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}(t) = \mathrm{Conv}(\mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}(r))$.
7. If t is the reflexivization of r, then $\mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}(t) = \mathrm{Ref}(\mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}(r))$.
8. If t is the predication$_k$ of r of s, then $\mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}(t) = \mathrm{Pred}_k(\mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}(r), \mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}(s))$.

This completes the semantics for L_ω.

15. A Complete Logic for the First Conception

On conception 1 intensional entities are identical if and only if necessarily equivalent. Thus, on conception 1 the following abbreviation captures the properties usually attributed to the modal operator $\Box$:

$$\Box A \ \textit{iff}_{\mathrm{df}} \ [A] = [[A] = [A]].$$

That is, necessarily A *iff* the proposition that A is identical to a trivial necessary truth. Since on conception 1 there is only one necessary truth, this definition is adequate. For the purpose of formulating the logic of L_ω on conception 1, this abbreviation will be adopted as a notational convenience. The modal operator $\Diamond$ is then defined in terms of $\Box$ in the usual way: $\Diamond A$ *iff*$_{\mathrm{df}}$ $\neg\Box\neg A$. By adopting these notational conventions, I am not reversing my earlier position on the parsing of natural language sentences such as 'it is necessary that A'. I would represent this sentence as $N([A])$. The 1-place predicate N may on conception 1 be defined as follows: $N(x)$ *iff*$_{\mathrm{df}}$ $x = [x = x]^x$.

The logic T1 for L_ω on conception 1 consists of the axiom schemas and rules for the modal logic S5 with quantifiers and identity and three additional axiom schemas for intensional abstracts.

Axiom Schemas and Rules of T1

A1: Truth-functional tautologies
A2: $(\forall v_i)A(v_i) \supset A(t)$ (where t is free for v_i in A)[11]
A3: $(\forall v_i)(A \supset B) \supset (A \supset (\forall v_i)B)$ (where v_i is not free in A)
A4: $v_i = v_i$

A5: $v_i = v_j \supset (A(v_i, v_i) \equiv A(v_i, v_j))$ (where $A(v_i, v_j)$ is a formula that arises from $A(v_i, v_i)$ by replacing some (but not necessarily all) free occurrences of v_i by v_j, and v_j is free for the occurrences of v_i that it replaces)

A6: $[A]_{u_1 \ldots u_p} \neq [B]_{v_1 \ldots v_q}$ (where $p \neq q$)

A7: $[A(u_1, \ldots, u_p)]_{u_1 \ldots u_p} = [A(v_1, \ldots, v_p)]_{v_1 \ldots v_p}$ (where these two terms are alphabetic variants)

A8: $[A]_\alpha = [B]_\alpha \equiv \Box(A \equiv_\alpha B)$

A9: $\Box A \supset A$

A10: $\Box(\mathrm{A} \supset \mathrm{B}) \supset (\Box\mathrm{A} \supset \Box\mathrm{B})$

A11: $\Box A \supset \Box\Diamond A$

R1: if $\vdash A$ and $\vdash (A \supset B)$, then $\vdash B$.

R2: if $\vdash A$, then $\vdash (\forall v_i)A$.

R3: if $\vdash A$, then $\vdash \Box A$.

A1 is, of course, concerned with the truth-functional sentential connectives & and $\neg$. A2 and A3 are familiar axioms for first-order quantifiers. A4 asserts the reflexivity of identity. A5 is Leibniz's law. A6 asserts the distinctness of intensional entities having different degrees. A7 asserts the validity$_1$ of a change of bound variables within intensional abstracts. A8 asserts the necessary equivalence of identicals and the identity of necessary equivalents. This principle is, of course, the hallmark of conception 1. A9–A11 are the standard S5 axioms for $\Box$ and $\Diamond$. R1 is modus ponens. R2 is universal generalization. R3 is the necessitation rule from S5.[12]

Given the definition $\Box$ and $\Diamond$ in terms of identity and intensional abstraction, modal logic may be viewed as the identity theory for intensional abstracts. In this connection, notice that, whereas the principle of necessary identity

$$x = y \supset \Box x = y$$

is an immediate consequence of Leibniz's law (A5) (given the reflexivity of identity (A4)), the S5 axiom (A11) is just an instance of the principle of necessary distinctness

$$x \neq y \supset \Box x \neq y.$$

In fact, the S5 axiom and the principle of necessary distinctness are actually equivalent. For, given A1–A10 and R1–R3, not only is A11 derivable from the principle of necessary distinctness, but also the principle of necessary distinctness is derivable from A11.

Now I will state the primary result for T1:

Theorem (Soundness and Completeness)
For all formulas A in L_ω, A is valid$_1$ if and only if A is a theorem of T1 (i.e., $\models_1 A$ iff $\vdash_{\mathrm{T1}} A$).

Proof (Soundness). First, the following lemmas are proved.
Lemma 1: T1 is equivalent to the theory that results when A5, A8, and A11 are replaced with the following simpler versions:

A5*	$v_i = v_j \supset (A(v_i, v_i) \supset A(v_i, v_j))$ (where $A(v_i, v_i)$ and $A(v_i, v_j)$ are as in A5 except that A is atomic)
A8*(a)	$\Box(A \equiv B) \equiv [A] = [B]$
A8*(b)	$(\forall v_i)([A(v_i)]_\alpha = [B(v_i)]_\alpha) \equiv [A(v_i)]_{\alpha v_i} = [B(v_i)]_{\alpha v_i}$
A11*	$v_i \neq v_j \supset \Box v_i \neq v_j$.

Lemma 2: Let v_h be an externally quantifiable variable in $[B(v_h)]_\alpha$, and let t_k be free for v_h in $[B(v_h)]_\alpha$. Consider any model structure $\mathscr{M}$ and any interpretation $\mathscr{I}$ and assignment $\mathscr{A}$ relative to $\mathscr{M}$. Let $\mathscr{A}'$ be an assignment that is just like $\mathscr{A}$ except that $\mathscr{A}'(v_h) = \mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}(t_k)$. Then, $\mathrm{D}_{\mathscr{I}\mathscr{A}'\mathscr{M}}([B(v_h)]_\alpha) = \mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}([B(t_k)]_\alpha)$.

Lemma 3: For all $\mathscr{I}$, $\mathscr{A}$, $\mathscr{M}$ and for all $\mathscr{D}_k \subset \mathscr{D} \in \mathscr{M}$, $k \geq 0$, $\mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}([A]_{v_1 \ldots v_k}) \in \mathscr{D}_k$.

Lemma 4: For all $\mathscr{I}$, $\mathscr{A}$, $\mathscr{M}$ and for all terms t and t', if $\mathscr{M}$ is type 1, then $\mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}([t = t]) = \mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}([t' = t'])$.

Lemma 5: Let v_r be free in $[A(v_r)]_\alpha$. Then, for all $\mathscr{I}$, $\mathscr{A}$, $\mathscr{M}$, if $\mathscr{M}$ is type 1, $\mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}([A(v_r)]_\alpha) = \mathrm{Pred}_0(\mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}([A(v_r)]_{\alpha v_r}), \mathscr{A}(v_r))$.

Lemma 6: For all $\mathscr{I}$, $\mathscr{A}$, $\mathscr{M}$,

(a) $\mathrm{T}_{\mathscr{I}\mathscr{A}\mathscr{M}}(F_i^j(t_1, \ldots, t_j))$ *iff* $\langle \mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}(t_1), \ldots, \mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}(t_j)\rangle \in \mathscr{G}(\mathscr{I}(F_i^j))$.
(b) $\mathrm{T}_{\mathscr{I}\mathscr{A}\mathscr{M}}((A \,\&\, B))$ *iff* $\mathrm{T}_{\mathscr{I}\mathscr{A}\mathscr{M}}(A)$ and $\mathrm{T}_{\mathscr{I}\mathscr{A}\mathscr{M}}(B)$.
(c) $\mathrm{T}_{\mathscr{I}\mathscr{A}\mathscr{M}}(\neg A)$ *iff* it is not the case that $\mathrm{T}_{\mathscr{I}\mathscr{A}\mathscr{M}}(A)$.
(d) $\mathrm{T}_{\mathscr{I}\mathscr{A}\mathscr{M}}((\exists v_k)A)$ *iff* there is an assignment $\mathscr{A}'$ relative to $\mathscr{M}$ such that $\mathscr{A}'$ is just like $\mathscr{A}$ except perhaps in what it assigns to v_k and $\mathrm{T}_{\mathscr{I}\mathscr{A}'\mathscr{M}}(A)$.

Then, given these lemmas, which are in most cases proofs by induction on the complexity of terms or formulas, the verification of the soundness of T1 is straightforward. (For example, the soundness of A6 follows directly from Lemma 3; the soundness of A8*(b), from Lemma 5, etc.)

Proof (Completeness). The proof is Henkin style. Let L^*_ω be any extension of L_ω. A sentence A is said to be derivable in T1 from set Γ of L^*_ω-sentences if, for some finite subset $\{B_1, \ldots, B_n\}$ of Γ, $\vdash_{T1} ((B_1 \,\&\ldots\&\, B_n) \supset A)$. A set $\mathscr{A}$ of sets of L^*_ω-sentences is said to be *perfect*$_1$ if (1) every set in $\mathscr{A}$ is maximal, consistent, and ω-complete; (2) for every identity sentence $t = t'$, if this sentence is in any set in $\mathscr{A}$, it is in all sets in $\mathscr{A}$; (3) for every sentence $[A]_{v_1 \ldots v_p} \neq [B]_{v_1 \ldots v_p}$ $(p \geq 0)$, if this sentence belongs to some $\Delta \in \mathscr{A}$, then there is some set $\Delta' \in \mathscr{A}$ (where possibly $\Delta = \Delta'$) such that the sentence $(\exists v_1) \ldots (\exists v_p) \neg (A \equiv B)$ belongs to Δ'; (4) for every closed term $[B]_{v_1 \ldots v_p}$, there is a primitive predicate letter F^p_q such that the sentence $[B]_{v_1 \ldots v_p} = [F^p_q(v_1, \ldots, v_p)]_{v_1 \ldots v_p} \in \Delta$, for some $\Delta \in \mathscr{A}$. The completeness of T1 follows from two lemmas.

Lemma 1: For every consistent set Γ of sentences in L_ω, there is a (denumerable) extension of L_ω relative to which there is a perfect$_1$ set $\mathscr{A}$ one of whose members Δ includes Γ.

Lemma 2: For every extension of L_ω relative to which $\mathscr{A}$ is a perfect$_1$ set, every set Δ in $\mathscr{A}$ has a type 1 model (whose cardinality is that of Δ).

To prove Lemma 1, I first form an extension L^*_ω of L_ω that has denumerably many primitive names and denumerably many i-ary primitive predicates for each $i \geq 0$. The sentences of L^*_ω are then arranged into a sequence of consecutive sentences $A_1, A_2, A_3, \ldots,$ having the following property: $A_1 = A_2$ and for every closed term $[B]_{v_1 \ldots v_p}$ in L^*_ω, there is at least one j such that A_j is the sentence $[B]_{v_1 \ldots v_p} = [F^p_q(v_1, \ldots, v_p)]_{v_1 \ldots v_p}$ where F^p_q is a primitive predicate letter that does not occur in B, Γ, or any A_h, $h < j$. Relative to this sequence, I use certain rules to construct an array of sets of L^*_ω-sentences:

$$\begin{array}{cccccc}
\Delta_1 & \Delta_3 & \Delta_7 & \cdots & \Delta_{n^2+n+1} & \cdots \\
\Delta_2 & \Delta_4 & \Delta_8 & & \Delta_{n^2+n+2} & \cdots \\
\Delta_5 & \Delta_6 & \Delta_9 & & \Delta_{n^2+n+3} & \cdots \\
\vdots & & \ddots & & \vdots & \\
 & & & \Delta_{n^2} & \Delta_{n^2+2n} & \cdots \\
\Delta_{n^2+1} & \Delta_{n^2+2} & \Delta_{n^2+3} \cdots & \Delta_{n^2+n} & \Delta_{(n+1)^2} & \cdots \\
\cdots\cdot & & & & &
\end{array}$$

The rules are these. (1) $\Delta_1 = \Gamma$. (2) If A_n, $n \geq 1$, is $[A]_\alpha \neq [B]_\alpha$ and $A_n \in \Delta_{n^2}$, then $\Delta_{n^2+1} = \{(\exists\alpha)\neg(A \equiv B)\}$; otherwise, $\Delta_{n^2+1} = \Delta_{n^2}$. (3) Let Δ_m, $m > 1$, be in column $i > 1$ and row $k \geq 1$. Then if $m^+ \cup m^* \cup \{A_i\}$ is consistent, $\Delta_m = m^+ \cup m' \cup \{A_i\}$; otherwise, $\Delta_m = m^+ \cup m'$. The sets m^+, m^*, and m' are:

$m^+ =_{\text{df}}$ the set in row k and column i–1
$m^* =_{\text{df}} \{[B]_\alpha = [C]_\beta : (\exists n < m)(\Delta_n \vdash_{\text{T1}} [B]_\alpha = [C]_\beta)\}$
$m' =_{\text{df}} \{C_1(a_1), \ldots, C_s(a_s)\}$

where the sentences $C_1(a_1), \ldots, C_s(a_s)$ are determined as follows: in the order in which they first occur in the sequence $A_1, A_2, \ldots, A_i, \ldots$, the sentences $(\exists v_1)C_1(v_1), \ldots, (\exists v_s)C_s(v_s)$ exhaust the existential sentences in m^+ that occur before A_i, and $C_1(a_1), \ldots, C_s(a_s)$ are the earliest substitution instances of $(\exists v_1)C_1(v_1), \ldots, (\exists v_s)C_s(v_s)$ occurring after A_i such that in order each $C_r(a_r)$, $1 \leq r \leq s$, contains the first occurence of the primitive name a_r anywhere in the sequence $A_1, A_2, \ldots, A_i, \ldots$. Now the set Δ^j is defined to be the union of all sets in row j, $j \geq 1$. And the set $\mathscr{A}$ is defined to be the set of all sets Δ^j, $j \geq 1$. Claim: $\mathscr{A}$ is perfect$_1$. This claim, which entails Lemma 1, can be proved once we have the following sublemma: for all $m \geq 1$, $\Delta_m \cup m^*$ is consistent. This sublemma is proved by induction on m.

Lemma 2 is proved as follows. Let L^*_ω be any extension of L_ω relative to which $\mathscr{A}$ is a perfect$_1$ set. For each $\Delta \in \mathscr{A}$, I construct a separate type 1 model $\langle \mathscr{M}_\Delta, \mathscr{I}_\Delta \rangle$. Choose some well-ordering $<$ of the union of the class of individual constants and the class of primitive predicate letters in L^*_ω, where $=$ is the least primitive predicate letter in this well-ordering. The domain $\mathscr{D}_\Delta$ is then identified with the following union:

$\{F_i^j \in \mathrm{L}_\omega^*$: there is no $F_h^k \in \mathrm{L}_\omega^*$ such that $F_h^k < F_i^j$ and the sentence $[F_h^k(v_1, \ldots, v_k)]_{v_1 \ldots v_k} = [F_i^j(u_1, \ldots, u_j)]_{u_1 \ldots u_j} \in \Delta\} \cup$
$\{a_j \in \mathrm{L}_\omega^*$: there is no $F_h^k \in \mathrm{L}_\omega^*$ such that the sentence $[F_h^k(v_1, \ldots, v_k)]_{v_1 \ldots v_k} = a_j \in \Delta$, and there is no $a_i \in \mathrm{L}_\omega^*$ such that $a_i < a_j$ such that the sentence $a_i = a_j \in \Delta\}$.

The subdomain $\mathscr{D}_{-1}$ is the set of primitive names in $\mathscr{D}_\Delta$, and the subdomain $\mathscr{D}_i$, $i \geq 0$, is the set of primitive i-ary predicates in $\mathscr{D}_\Delta$. The prelinear ordering $\mathscr{P}$ is defined as follows: $\mathscr{P}(x, y)$ *iff*$_{\mathrm{df}}$ for some i and j, $i < j$, $x \in \mathscr{D}_i$ and $y \in \mathscr{D}_j$. The set $\mathscr{K}$ of alternate extension functions $H_{\Delta'}$, is determined by the atomic sentences belonging to the various sets Δ' belonging to $\mathscr{A}$. The actual extension function $\mathscr{G} =_{\mathrm{df}} H_\Delta$. The identity element $\mathrm{Id} \in \mathscr{M}_\Delta$ is just the identity predicate $=$. And the fundamental logical operations Conj_Δ, Neg_Δ, $\mathrm{Exist}_\Delta, \ldots$ are determined by the identity sentences in Δ. Finally, the interpretation $\mathscr{I}_\Delta$ may be defined as follows:

$\mathscr{I}_\Delta(`a_i\text{'}) =_{\mathrm{df}}$ the individual constant $`a_j\text{'} \in \mathscr{D}$ such that $`a_i = a_j\text{'} \in \Delta$
$\mathscr{I}_\Delta(`F_i^j\text{'}) =_{\mathrm{df}}$ the primitive predicate $`F_k^j\text{'} \in \mathscr{D}$ such that $`[F_i^j(v_1, \ldots, v_j)]_{v_1 \ldots v_j} = [F_k^j(v_1, \ldots, v_j)]_{v_1 \ldots v_j}\text{'} \in \Delta$.

With $\mathscr{M}_\Delta$ and $\mathscr{I}_\Delta$ so specified, it is then shown by induction on the complexity of formulas that $\langle \mathscr{M}_\Delta, \mathscr{I}_\Delta \rangle$ is a model of Δ, for all $\Delta \in \mathscr{A}$.

By the way, the completeness theorem for T1 yields an interesting corollary. Notice that L_ω is a notational variant of a first-order extensional language that is fitted out with identity and extensional abstracts $\{v_1 \ldots v_j: A\}$, for $j \geq 0$. Let an extensional type 1 model structure be defined to be a type 1 model structure in which the class $\mathscr{K}$ of alternate extension functions is just $\{\mathscr{G}\}$ (i.e., the singleton of the actual extension function). Thus, in an extensional type 1 model structure the following holds for all $i \geq -1$:

$$(\forall x, y \in \mathscr{D}_i)(\mathscr{G}(x) = \mathscr{G}(y) \supset x = y).$$

And hence, the elements of $\mathscr{D}$ behave as extensional entities do. The semantics is done in precisely the same way in which the semantics for L_ω is done except that only extensional type 1 model structures are considered. This yields the notion of *extensional validity*. And the formal logic consists of the axiom schemas and rules for standard first-order quantifier logic with identity (i.e., A1–A5, R1–R2) plus three axioms schemas for extensional abstracts:

(i) $\{u_1 \ldots u_p\colon A\} \neq \{v_1 \ldots v_q\colon B\}$ (where $p \neq q$)
(ii) $\{u_1 \ldots u_p\colon A(u_1, \ldots, u_p)\} = \{v_1 \ldots v_p\colon A(v_1, \ldots, v_p)\}$
(where the externally quantifiable variables in these two complex terms are the same and, for each k, $1 \leq k \leq p$, u_k is free in A for v_k and conversely)
(iii) $(A \equiv_{v_1 \ldots v_p} B) \equiv \{v_1 \ldots v_p\colon A\} = \{v_1 \ldots v_p\colon B\}$.

Schema (i) asserts the distinctness of truth values from sets and relations-in-extension, the distinctness of sets from relations-in-extension, and the distinctness of m-ary relations-in-extension from n-ary relations-in-extension ($m \neq n$). Schema (ii) asserts the validity of a change of bound variables within extensional abstracts. And schema (iii) asserts the equivalence of identicals and the identity of equivalents. This property is the hallmark of extensional entities. The primary result for this extensional logic is the following corollary of the completeness and soundness theory for T1:

Corollary (Soundness and Completeness)
For all formulas A in a first-order extensional language with identity and extensional abstraction, A is extensionally valid if and only if A is a theorem of the logic for the language.

Thus, when $\in$ is treated as an arbitrary first-order predicate, set theory with identity and extensional abstraction is sound and complete.

16. A Complete Logic for the Second Conception

On conception 2 each definable intensional entity is such that when it is defined completely, it has a unique, non-circular definition. The logic T2 for L_ω on conception 2 consists of axioms A1–A7 and rules R1–R2 from T1, five additional axiom schemas for intensional abstracts, and one additional rule. In stating the additional principles, I write $t(F^n_m)$ to indicate that t is a complex term of L_ω in which the primitive predicate F^n_m occurs.

Additional Axiom Schemas and Rules for T2

$\mathscr{A}8$: $[A]_\alpha = [B]_\alpha \supset (A \equiv B)$

$\mathscr{A}9$: $t \neq r$ (where t and r are non-elementary complex terms of different syntactic kinds)

$\mathscr{A}10$: $t = r \equiv t' = r'$ (where t and r are the negations (existential generalizations, expansions, inversions, conversions, reflexivizations) of t' and r', respectively)

$\mathscr{A}11$: $t = r \equiv (t' = r' \;\&\; t'' = r'')$ (where t is the conjunction of t' and t'' and r is the conjunction of r' and r'', or t is the predication$_k$ of t' of t'' and r is the predication$_k$ of r' of r'', for $k \geq 0$)

$\mathscr{A}12$: $t(F^j_i) = r(F^k_h) \supset q(F^j_i) \neq s(F^k_h)$ (where t and s are elementary and r and q are not)

$\mathscr{R}3$: Let F^n_m be a non-logical predicate that does not occur in $A(v_i)$; let $t(F^n_m)$ be an elementary complex term, and let t' be any complex term of degree n that is free for v_i in $A(v_i)$. If $\vdash A(t)$, then $\vdash A(t')$.

$\mathscr{A}8$ affirms the equivalence of identical intensional entities. Schemas $\mathscr{A}9$–$\mathscr{A}11$ capture the principle that a complete definition of an intensional entity is unique. And schema $\mathscr{A}12$ captures the principle that a definition of an intensional entity must be non-circular. $\mathscr{R}3$ says roughly that if $A(t)$ is valid$_2$ for an arbitrary elementary n-ary term t, then $A(t')$ is valid$_2$ for any n-ary term t'.

Now recall the two intuitively valid arguments mentioned at the outset of this chapter. As we have seen, these arguments may be symbolized in L_ω as follows:

$$\begin{array}{l} (\forall z)(B(x, z) \supset B(y, z)) \\ B(x, [A]) \\ \hline \therefore\ B(y, [A]) \end{array}$$

$$\begin{array}{l} [B(x)]_x = [U(x) \;\&\; M(x)]_x \\ N([(\forall x)(B(x) \equiv B(x))]) \\ \hline \therefore\ N([(\forall x)(B(x) \equiv (U(x) \;\&\; M(x)))]). \end{array}$$

These arguments are both valid$_1$ and valid$_2$, and relatedly, in both T1 and T2 the conclusion of each argument is derivable from its premises.

To bring out the difference between T1 and T2 (and between validity$_1$ and validity$_2$), consider the following intuitively *invalid* argument involving the intentional predicate 'wonders':

x wonders whether there is a trilateral that is not a triangle.
Necessarily, all and only trilaterals are triangles.

∴ x wonders whether there is a triangle that is not a triangle.

Let this argument be symbolized as follows:

$$xW[(\exists y)(\text{Trilateral }(y)\ \&\ \neg\text{Triangle }(y))]$$
$$\Box(\forall y)(\text{Trilateral }(y) \equiv \text{Triangle }(y))$$
$$\therefore\ xW[(\exists y)(\text{Triangle }(y)\ \&\ \neg\text{Triangle }(y))].$$

In T1, but not T2, the conclusion of this argument is derivable from the two premises. And relatedly, the argument is valid$_1$, but not valid$_2$. So only the formal logic and semantics that are based on conception 2 could be appropriate for the treatment of intentional matters. The fact that Church's Alternative (2) and the various possible-worlds constructions of intensional logic (including Carnap's original construction in *Meaning and Necessity*) are all based on conception 1 is what lies at the root of their failure to provide adequate treatments of intentional matters. (See desideratum 2 on the chart in §4.)

The following is the primary result for T2:

Theorem (Soundness and Completeness)
For all formulas A in L_ω, A is valid$_2$ if and only if A is a theorem of T2 (i.e., $\models_2 A$ iff $\vdash_{\mathrm{T2}} A$).

Proof. The proof of the soundness of T2 is quite straightforward. For example, the soundness of $\mathscr{A}8$ follows directly from Lemma 6 (stated earlier); $\mathscr{A}9$, from the fact that the fundamental logical operations Conj, Neg, Exist, ... in a type 2 model structure have disjoint ranges; $\mathscr{A}10$ and $\mathscr{A}11$, from the fact that these functions are one-one; $\mathscr{A}12$, from the fact that they are non-cycling. The soundness proofs for R1 and R2 are standard. For the soundness of $\mathscr{R}3$, the induction hypothesis yields $\models_2 A(t(F^n_m))$. Hence, by the soundness of R2, A2, and A5 (Leibniz's law), we have $\models_2 t(F^n_m) = t' \supset A(t')$. But since F^n_m is a non-logical predicate and does not occur in $A(t')$, $\models_2 A(t')$. The completeness proof is again

Henkin style. A set of L_ω^*-sentences is said to be *perfect*$_2$ if (1) it is maximal, consistent, ω-complete and (2) for every closed term $[B]_{v_1 \ldots v_p}$ in L_ω^*, there is a primitive predicate letter F_k^p such that the sentence $[B]_{v_1 \ldots v_p} = [F_k^p(v_1, \ldots, v_p)]_{v_1 \ldots v_p}$ belongs to the set. I show, first, that every consistent set of L_ω-sentences is included in some perfect$_2$ set of L_ω^*-sentences and, secondly, that every perfect$_2$ set has a type 2 model. The argument, while parallel to the argument used for T1, is routine.

The completeness problem for T2 is essentially simpler than the one for T1. This is no reflection on the relative importance of T2, though, for T2, not T1, provides a logic for intentional matters.

17. A Complete Logic for Modal and Intentional Matters

The conception 1 intensional logic T1 is ideally suited for treating modal matters. And the conception 2 intensional logic T2 is ideally suited for treating intentional matters. I will now formulate a richer conception 2 logic T2′ that is ideally suited for treating both modal and intentional matters. This simultaneous treatment is achieved by adjoining to L_ω a 2-place logical predicate $\approx_N$ which is intended to express the relation of necessary equivalence. T2′ succeeds in providing a single logic for both modal and intentional matters by having what are in effect two sorts of "identity"—one weak and one strong. The former is necessary equivalence; the latter, strict identity. In §46 I will show that, when conceptions 1 and 2 are synthesized, necessary equivalence (and also necessity) can be defined. So we should not feel hesitant to adjoin $\approx_N$ to L_ω here.

I begin by defining a new type of model structure. A *type 2′* model structure

$$\langle \mathscr{D}, \mathscr{P}, \mathscr{K}, \mathscr{G}, \mathrm{Id}, \mathrm{Eq_N}, \mathrm{Conj}, \mathrm{Neg}, \mathrm{Exist}, \\ \mathrm{Exp}, \mathrm{Inv}, \mathrm{Conv}, \mathrm{Ref}, \mathrm{Pred}_0, \mathrm{Pred}_1, \ldots \rangle$$

is any structure that satisfies all the conditions imposed on type 2 model structures plus one additional condition. The element $\mathrm{Eq_N}$ must be a distinguished element of $\mathscr{D}_2$ that satisfies the following principle:

$$(\forall H \in \mathscr{K})(H(\mathrm{Eq_N}) = \\ \{xy\colon (\exists i \geq -1)x, y \in \mathscr{D}_i \ \& \ (\forall H' \in \mathscr{K})H'(x) = H'(y)\}).$$

Thus Eq_{N} is to be thought of as the distinguished logical relation-in-intension *necessary equivalence*. Now an interpretation $\mathscr{I}$ relative to a type 2′ model structure is just like an interpretation relative to a type 1 or type 2 model structure except that $\mathscr{I}(\approx_{\mathrm{N}}) = \mathrm{Eq}_{\mathrm{N}}$. Then type 2′ denotation, truth, and validity are defined *mutatis mutandis* as in §14. The following abbreviations are introduced for notational convenience:

$\Box A$ *iff*$_{\mathrm{df}}$ $[A] \approx_{\mathrm{N}} [[A] \approx_{\mathrm{N}} [A]]$
$\Diamond A$ *iff*$_{\mathrm{df}}$ $\neg\Box\neg A$.

The intensional logic T2′ simply consists of the axioms and rules for T2 plus the following additional axioms and rules for $\approx_{\mathrm{N}}$:

$\mathscr{A}13$: $x \approx_{\mathrm{N}} x$
$\mathscr{A}14$: $x \approx_{\mathrm{N}} y \supset y \approx_{\mathrm{N}} x$
$\mathscr{A}15$: $x \approx_{\mathrm{N}} y \supset (y \approx_{\mathrm{N}} z \supset x \approx_{\mathrm{N}} z)$
$\mathscr{A}16$: $x \approx_{\mathrm{N}} y \supset \Box x \approx_{\mathrm{N}} y$
$\mathscr{A}17$: $\Box(A \equiv_{\alpha} B) \equiv [A]_{\alpha} \approx_{\mathrm{N}} [B]_{\alpha}$
$\mathscr{A}18$: $\Box A \supset A$
$\mathscr{A}19$: $\Box(A \supset B) \supset (\Box A \supset \Box B)$
$\mathscr{A}20$: $\Box A \supset \Box\Diamond A$
$\mathscr{R}4$: if $\vdash A$, then $\vdash \Box A$.

Notice that these axioms and rules for $\approx_{\mathrm{N}}$ are just analogues of the special T1 axioms and rules for $=$. Finally, the soundness and completeness of T2′ can be shown by applying the methods of proof used for T1 and T2.

Intensional logic constitutes the first stage in the theory of PRPs. Why is it that complete intensional logics can be achieved in the setting of a first-order language such as L_{ω} but not in the setting of a higher-order language? The answer lies in the opposing treatments of *predication*. To this, the second stage in the theory of PRPs, I will soon turn. But first I must address a problem that is perhaps the major outstanding problem in intensional logic, namely, the paradox of analysis.

3

The Paradox of Analysis*

18. The Paradox

The paradox of analysis is an important and complex problem in the philosophy of logic and language. What makes it important are its deep implications for philosophy in the areas of philosophical methodology and philosophical psychology and for psychology in the areas of development, perception, decision, and perhaps psychoanalysis. Yet in recent years philosophers have all but forgotten the problem.

When I speak of the paradox of analysis I am referring to logical puzzles of the following sort. Take three formulas:

(1) *x* knows that whatever is a circle is a circle.
(2) *x* does not know that whatever is a circle is a locus of points in the same plane equidistant from some common point.
(3) Being a circle = being a locus of points in the same plane equidistant from some common point.

The paradox is this: (1) and (2) are simultaneously satisfiable; (3) is true, and yet the conjunction of (1) and (3) entails the negation of (2). Hence, a contradiction.[1]

To get the paradox of analysis squarely in mind, one must distinguish it from a superficially similar yet fundamentally different problem which I call *Mates' puzzle*.[2] Mates' puzzle is generated in an analogous way by formulas such as the following:

(4) *x* knows that whatever chews chews.
(5) *x* does not know that whatever masticates chews.
(6) Being something that masticates = being something that chews.

* The reader may skip this chapter without interrupting the larger line of development in the book.

It would seem that (4) and (5) can be satisfied at the same time, for even though x knows what chewing is, he can in some sense fail to know what masticating is. Such a situation might arise as follows. Our person x has the concept of chewing (and *eo ipso* the concept of masticating). In addition, x knows that the predicate 'chew' expresses this concept. However, x does not know what the predicate 'masticate' expresses. Indeed, x's contact with the historical information chain associated with the English predicate 'masticate' does not go beyond his knowledge that English speakers have such a chain. x's ignorance as reported in (5) originates in his ignorance of linguistic (or historical or social) matters.[3] In the case of a genuine paradox of analysis, by contrast, the sort of ignorance at work does not originate in linguistic (or historical or social) ignorance. In the above instance of the paradox of analysis, for example, we can imagine that, besides being fully aware of what circularity is, x also knows that the English word 'circle' expresses circularity. What x is ignorant of is neither the concept of circularity nor the semantics of his language nor the relation between historical information chains. Rather, x is ignorant of the *definition* of circularity itself. Herein lies the difference between the genuine paradox of analysis and Mates' puzzle. The challenge posed by the paradox of analysis then is to find a satisfactory way to represent the non-linguistic (non-historical, non-social) knowledge that one acquires when one learns a definition. This is no trvial affair.

A number of people have proposed informal resolutions to the paradox of analysis. But in my view none of these informal resolutions promises to be adequate. In the field of formal intensional logic the paradox of analysis has been virtually ignored. Alonzo Church is the only logician to have incorporated into a formal theory of propositions a serious attempt to resolve the paradox.[4] And so it is to this resolution that I now turn.

19. Difficulties in Church's Resolution

I begin this section by sketching Church's resolution. Then I will show why it too is inadequate. One purpose for engaging in this exercise is to gain a better understanding of what is required of a resolution of the paradox.

To understand Church's resolution of the paradox of analysis, one must be familiar with his theory of synonymy, which he first states in 'Intensional Isomorphism and the Identity of Belief' (pp. 66–7).[5] On Church's theory, expressions are synonymous if and only if they are synonymous isomorphic. And expressions are *synonymous isomorphic* iff_{df} one can be obtained from the other by a series of steps that consists of (1) alphabetic changes of a bound variable, (2) replacement of one individual constant by another that is synonymous with it, (3) replacement of one predicate constant by another that is synonymous to it, (4) replacement of an abstraction expression—i.e., an expression of the form $(\lambda x)(\ldots x \ldots)$—by a synonymous predicate constant, (5) replacement of a predicate constant by a synonymous abstraction expression, (6) replacement of an individual description by a synonymous individual constant, (7) replacement of an individual constant by a synonymous individual description.

If correct, Church's theory of synonymy yields a resolution of "one half" of the instances of the paradox of analysis. To see how this works, consider the instance of the paradox of analysis generated in the usual way by the following formulas:

(7) x knows that, for all y and z, if $y = z$, then $y = z$.
(8) x does not know that, for all y and z, if every property of y is a property of z and conversely, then $y = z$.
(9) Being y and z such that $y = z$ is the same as being y and z such that every property of y is a property of z and conversely.[6]

Church would resolve this instance of the paradox by denying (9). Since '$y = z$' cannot be converted by rules (1)–(7) into 'Every property of y is a property of z and conversely', they are not synonymous isomorphic.[7] So on Church's theory of synonymy they are not synonymous, and therefore, the concepts they express are different. Thus, (9) is false, and the contradiction is avoided.

But consider the following formulas adapted from an example discussed at length by Church:[8]

(10) x knows that whatever is a period lasting fourteen days is a period lasting fourteen days.
(11) x does not know whatever is a fortnight is a period lasting fourteen days.
(12) Being a fortnight = being a period lasting fourteen days.

These three formulas give rise to a contradiction in the usual way. Since 'x is a fortnight' can be converted into 'x is a period lasting fourteen days' by rule (5), these two expressions are synonymous isomorphic. Thus, on Church's theory of synonymy they are synonymous. So they express the same concept, and hence (12) is true. At the same time, given Church's axioms for the logic of sense and denotation,[9] the negation of (11) is derivable from the conjunction of (10) and (12). Thus, Church must deny that (10) and (11) are simultaneously satisfiable. This is what he does. But since we do have an intuition that (10) and (11) are simultaneously satisfiable, an explanation of why it is mistaken ought to be given. Church's explanation goes as follows.[10] Suppose that (10) is satisfied. Then the possibility that we take to be expressed by (11) is not in fact expressed by (11). Rather, it is expressed by the following formula containing metalinguistic terminology:

(11′) x does not know that whatever satisfies the English sentential matrix 'y is a fortnight' satisfies the English sentential matrix 'y is a period lasting fourteen days'.

Clearly, (10) and (11′) are simultaneously satisfiable. Indeed, (10), (11′), and (12) are jointly consistent. Extrapolating from the foregoing, I then arrive at the following conclusion. If there are instances of the paradox of analysis whose key expressions are synonymous isomorphic, Church is committed to resolving these instances by means of the above metalinguistic maneuver. We shall see below that there indeed are such instances of the paradox. And so it is that Church is led to resolve the "second half" of the instances of the paradox by means of the metalinguistic maneuver.

Now I turn to the assessment of Church's two types of resolutions of the paradox of analysis. The first type of resolution is inadequate because the theory of synonymy is mistaken. This can be shown in two ways. The first is simply by example. Intuitively, 'y outweighs z' and 'The weight of y is greater than the weight of z' are synonymous; however, on Church's theory they could not be since they are not synonymous isomorphic.[11] This is not just an isolated counterexample; examples of this sort abound.[12] The second way to show the inadequacy of Church's theory is this. By radically limiting the synonym pairs, the theory acts as a two-edged sword: although in its way it does resolve certain instances of the paradox of analysis, it has the undesirable effect of artificially shrinking the logical-consequence relation for propositions.[13] To see this, consider the following two formulas:

> x believes that, if the weights of y and z differ and if y does not outweigh z, then z outweighs y.
>
> x believes that, if the weights of y and z differ and if y does not outweigh z, then the weight of z is greater than the weight of y.

There are readings of these formulas (perhaps the most natural readings) according to which the proposition expressed by the latter formula is a logical consequence of the proposition expressed by the former formula. However, Church's theory of synonymy makes these readings impossible. For this and the preceding reason I conclude that Church's theory is mistaken. After all, the only motivation for Church's theory appears to be that in its way it can help to resolve the paradox of analysis.

The major flaw in Church's second type of resolution, his metalinguistic resolution, is that it in effect rules out the possibility of *informative definitions* beyond those that concern mere linguistic facts. That is, it is a consequence of the metalinguistic resolution that what one learns when one discovers a correct definition is merely something about a language. Yet in view of Church's famous criticism of Carnap's analysis of statements of assertion and belief, it is surprising that Church should take this line, for a criticism akin to Church's criticism of Carnap can now be lodged against Church. To illustrate this criticism, let us consider the instance of the paradox of analysis given at the outset of the chapter. On analogy with 'y is a fortnight' and 'y is a period lasting fourteen days', the expressions 'y is a circle' and 'y is a locus of points in the same plane equidistant from some common point' are synonymous isomorphic. Hence, Church would hold that (3) is true. Therefore, on analogy with his treatment of the earlier 'fortnight' case, Church would be led to deny that (1) and (2) are simultaneously satisfiable. And in turn he would be led to hold that, if (1) is satisfied, then the possibility that we mistakenly take to be expressed by (2) is actually expressed by:

(2′) x does not know that whatever satisfies the English sentential matrix 'y is a circle' satisfies the English sentential matrix 'y is a locus of points in the same plane equidistant from some common point'.

Since (1), (2′), and (3) are jointly consistent, Church would hold that the paradox is resolved. However, let us consider a deaf-mute x

whom we know to know no English (and, we may suppose, no other language either). Suppose that we observe x sorting out circular objects from non-circular ones. On the basis of this and related evidence we infer that x has the concept of circularity. In turn, we infer the proposition expressed by (1). Now suppose I ask you, 'Does x know that whatever is a circle is a locus of points in the same plane equidistant from some common point?'. Shortly thereafter we observe x performing a variety of relevant geometric constructions with ruler and compass. On the basis of these observations, I make the inference that prompts me to assert, 'x does know that whatever is a circle is a locus of points in the same plane equidistant from some common point'. Now according to Church's view, I have asserted the proposition expressed by the following metalinguistic formula:

> x knows that whatever satisfies the English sentential matrix 'y is a circle' satisfies the English sentential matrix 'y is a locus of points in the same plane equidistant from some common point'.

But this does not seem credible. I have not made an assertion about x's knowledge of English; indeed, I know that x knows no English. Nor have I (*à la* Carnap) made an assertion about x's knowledge of some other language somehow related to English; we may suppose that x knows no language at all. When people (regardless of which language, if any, they speak) learn that circles are loci of points in the same plane equidistant from some common point, they learn a fact about circles, not about language.[14] Here again we see that ignorance of conceptual definitions (analyses) is not a species of linguistic ignorance.

20. A New Resolution

The foregoing difficulties in Church's resolution suggest, first, that the paradox cannot be resolved by imposing more tight-fisted criteria of synonymy and, secondly, that the ignorance at the heart of the paradox is not linguistic in nature. This leads one to think that the paradox results instead from an *ambiguity* in intensional abstracts in natural language. On one way of reading intensional abstracts, substitutivity of synonyms is guaranteed; on another it is violated. The reading on which substitutivity is violated is that which is associated with ignorance of definitions, the source of the paradox. Consider the following intensional abstract:

the proposition that whatever is a circle is a locus of points in the same plane equidistant from some common point.

There is a reading of this abstract (reading I) according to which it denotes a different proposition from the one we would usually take to be denoted by:

the proposition that whatever is a circle is a circle.

And there is another reading (reading II) according to which it denotes the same proposition. Now we saw earlier that, considered one at a time, the following claims about sentences (1)–(3) seemed intuitively to hold:

(a) (1) and (2) are simultaneously satisfiable.
(b) (3) is true.
(c) The conjunction of (1) and (3) entails the negation of (2).

This leads to paradox since (a), (b), and (c) are jointly inconsistent. This paradox vanishes, though, when attention is paid to the two possible readings of the intensional abstract in (2). When it is given reading I, formulas (1) and (2) are simultaneously satisfiable. But the conjunction of (1) and (3) entails the negation of (2) only when the abstract in (2) is given reading II. The paradox thus is only a fallacy of equivocation.

It would be inelegant, perhaps even impossible, to treat each ambiguity of this sort as a semantic ambiguity, i.e., as an ambiguity resulting from a plurality of meanings of primitive non-logical constants. It is preferable to treat it as a structural (viz., syntactic) ambiguity, i.e., an ambiguity resulting from a plurality of syntactic deep structures that lead to intensional abstracts having the same surface syntactic structure.[15] This can be accomplished simply by introducing to L_ω a new syntactic operation according to which any open or closed sentence within an intensional abstract may be *underlined*. For heuristic purposes we may think of an underline as indicating that the intension expressed by an underlined open sentence is an *undefined concept* and the intension expressed by an underlined closed sentence is an *undefined thought*. (In these preliminary remarks it might be helpful to think of undefined concepts and thoughts as type 1 intensions, and defined concepts and thoughts as type 2 intensions; we shall see later that there is at least one further interesting way to conceive of defined and undefined concepts and thoughts.)

To see how this underlining device works, consider again the previous instance of the paradox that concerns identity. Formulas (7)–(9) are on one reading inconsistent with one another. This reading is now represented as follows:*

$$(7')\quad xK[(\forall y, z)(\underline{y = z} \supset \underline{y = z})]$$

$$(8')\quad \neg xK[(\forall y, z)(\underline{(\forall w)(y \,\Delta\, w \equiv z \,\Delta\, w)} \supset \underline{y = z})]$$

$$(9')\quad [\underline{y = z}]_{yz} = [\underline{(\forall w)(y \,\Delta\, w \equiv z \,\Delta\, w)}]_{yz}.$$

The inconsistency arises since (9′) entails

$$[(\forall y, z)(\underline{y = z} \supset \underline{y = z})] = [(\forall y, z)(\underline{(\forall w)(y \,\Delta\, w \equiv z \,\Delta\, w)} \supset \underline{y = z})]$$

which, together with (7′), contradicts (8′). Before, we seemed to have a paradox on our hands because only an inconsistent reading of (7)–(9) could be represented; yet intuitively (7)–(9) were consistent. We can now avoid the paradox, though, for we can now represent the consistent reading by replacing (8′) with:

$$(8'')\quad \neg xK[(\forall y, z)((\forall w)(y \,\Delta\, w \equiv z \,\Delta\, w) \supset \underline{y = z})].$$

The only difference between (8″) and (8′) is of course that $(\forall w)(y \,\Delta\, w \equiv z \,\Delta\, w)$ is not underlined in (8″) whereas it is in (8′). Since $[\underline{(\forall w)(y \,\Delta\, w \equiv z \,\Delta\, w)}]_{yz}$ denotes the undefined identity concept and since $[(\forall w)(y \,\Delta\, w \equiv z \,\Delta\, w)]_{yz}$ denotes a defined identity concept, we know that

$$[\underline{(\forall w)(y \,\Delta\, w \equiv z \,\Delta\, w)}]_{yz} \neq [(\forall w)(y \,\Delta\, w \equiv z \,\Delta\, w)]_{yz}$$

and, hence, that

$$[(\forall y, z)(\underline{y = z} \supset \underline{y = z})] \neq [(\forall y, z)((\forall w)(y \,\Delta\, w \equiv y \,\Delta\, z) \supset \underline{y = z})].$$

This is what makes (7′), (8″), and (9′) consistent with one another, as the most natural reading of (7), (8), and (9) calls for. Thus, by enriching L_ω syntactically we successfully avoid the paradox.

It is easy to construct this enriched language, which will be called $\underline{L}_\omega$. The primitive symbols of $\underline{L}_\omega$ are those of L_ω plus the underline. The simultaneous inductive definition of term and formula for $\underline{L}_\omega$ goes as follows:

* Note Δ is a distinguished 2-place logical predicate that expresses the predication relation. For more on Δ, see the next chapter. Incidentally, there is an alternate inconsistent reading of (7)–(9) which can be represented by formulas that are like (7′)–(9′) except that all underlines are omitted.

(1) All variables are terms.
(2) If $t_1, \ldots, t_j$ are terms, then $F_i^j(t_1, \ldots, t_j)$ is a formula.
(3) If A and B are formulas and v_k is a variable, then $(A \,\&\, B)$, $\neg A$, and $(\exists v_i)A$ are formulas.
(4) If the expression A' is just like a formula A except perhaps that some subformulas occurring in A are underlined in A' and if $v_1, \ldots, v_p$ (for $p \geq 0$) are distinct variables, then $[A']_{v_1 \ldots v_p}$ is a term.
(5) If $[\underline{A}]_{v_1 \ldots v_p}$ (for $p \geq 0$) is a term, then so is $\underline{[A]_{v_1 \ldots v_p}}$.[16]

The indicated structural ambiguities in intensional abstracts in natural language can be unambiguously represented in $\underline{L}_\omega$. Since these structural ambiguities are responsible for the paradox of analysis, the paradox should not arise in $\underline{L}_\omega$. However, before this can be guaranteed, the right type of semantics for $\underline{L}_\omega$ must be specified.

Since the aim is simply to characterize the logically valid formulas of $\underline{L}_\omega$, a Tarski-style semantics will suffice. And what we are lacking at present is simply a method for modeling the heuristic distinction between undefined and defined ideas. There are two intriguing candidate methods, each of which calls for further philosophical study. The first invokes the theory of qualities and concepts which is developed in chapter 8. Specifically, undefined concepts would be identified with qualities (or connections), which are type 1 intensions, and defined concepts would be identified with complex concepts, which are type 2 intensions. This method is especially appealing since it meshes so nicely with the Platonic account of genuine forms and their analysis. The second method would just posit outright two primitive sorts of type 2 intensions, one corresponding to undefined concepts and the other to defined concepts. Either way, the semantics for $\underline{\underline{L}}_\omega$ is a straightforward affair once the appropriate types of model structures have been specified.[17] Relative to such semantics, it is easy to formulate logics for $\underline{\underline{L}}_\omega$ which can, it appears, be proven both sound and complete. Given this, a full resolution of the paradox of analysis is at hand.[18]

This finishes my study of intensional logic, the first stage in the study of PRPs. This complete foundation readies us for the second stage, the extension of intensional logic to include the predication relation.

Part II

Extension

4

Predication

21. The First-Order/Higher-Order Controversy

First-order quantifier logic is complete; higher-order quantifier logic is not. A few formally minded philosophers of logic—such as Quine and some of his followers—appear to believe that this is sufficient grounds for concluding that the only legitimate quantifier logic is first-order, not higher-order. However, most leading formally minded philosophers of logic over the past hundred years—Frege, Russell, Church, Carnap, Henkin, Montague, Kaplan—believe that the higher-order approach is a natural generalization of the first-order approach and therefore that quantifier logic is properly identified with higher-order quantifier logic. I depart from this majority opinion. In §10 I gave several formalistic reasons (including completeness) for preferring the first-order approach over the higher-order approach. But formalistic reasons tell us little about the issues of naturalness and generality. In this chapter I will discuss the underlying philosophical differences between the two approaches to quantifier logic.[1] My hope is that the greater naturalness and generality of the first-order approach will become evident in the course of the discussion.

Consider the following intuitively valid argument:

x is red and y is not red.

∴ There is something that x is and that y is not.

There are two approaches to the representation of this argument—the first-order approach and the higher-order approach. On the higher-order approach the argument is represented as an instance of second-order existential generalization:

$$\frac{Rx \,\&\, \neg Ry}{\therefore\ (\exists f)(fx \,\&\, \neg fy)}$$

where R is a name of the color red and f is a predicate variable for which R is a substituend. On the first-order approach the argument is represented as an instance of first-order existential generalization:

$$\begin{array}{l} \quad x \,\Delta\, r \,\&\, y \not\Delta\, r \\ \hline \therefore\ (\exists z)(x \,\Delta\, z \,\&\, y \not\Delta\, z) \end{array}$$

where r is a name of the color red and Δ is a distinguished 2-place logical predicate that expresses the *predication relation*, a relation expressed by the *copula* in natural language.[2]

There are analogous examples involving relations rather than properties:

$$\begin{array}{l} \quad \text{x and y are husband and wife, and u and v are not husband} \\ \quad \text{and wife.} \\ \hline \therefore\ \text{There is something which x and y are that u and v are not.} \end{array}$$

On the higher-order approach this intuitively valid argument is represented as an instance of second-order existential generalization:

$$\begin{array}{l} \quad H^2(x, y) \,\&\, \neg H^2(u, v) \\ \hline \therefore\ (\exists f^2)(f^2(x, y) \,\&\, \neg f^2(u, v)) \end{array}$$

where the 2-place predicate H^2 is construed as a name for the relation holding between husband and wife and f^2 is a 2-place predicate variable. On the first-order approach the argument is represented as an instance of first-order existential generalization:[3]

$$\begin{array}{l} \quad \langle x, y\rangle \,\Delta\, [H^2(x, y)]_{xy} \,\&\, \langle u, v\rangle \not\Delta\, [H^2(x, y)]_{xy} \\ \hline \therefore\ (\exists z)(\langle x, y\rangle \,\Delta\, z \,\&\, \langle u, v\rangle \not\Delta\, z). \end{array}$$

Philosophically speaking, how do the higher-order and first-order approaches differ? In the next few sections I suggest an answer to this question.

22. Expressive Power

It is often thought that a higher-order language has greater expressive power than the first-order counterpart. However, for appropriate first-order languages (such as L_ω with Δ and $\approx_N$), this is not so, and indeed the situation is typically the other way around. In fact, since the variables in L_ω are free to range over all the objects falling within the range of any given higher-order variable, higher-order notation can be contextually defined in L_ω with Δ and $\approx_N$.

For illustrative purposes I will show how this can be done for a sample second-order language. More complex higher-order languages can be dealt with on analogy.[4] I begin with some preliminary definitions:

$$\text{Individual Particular } (x) \ \mathit{iff}_{\mathrm{df}} \ \Box(\forall y)(y \,\Delta\, x \equiv y = x)$$

$$\text{Proposition } (x) \ \mathit{iff}_{\mathrm{df}} \ (\exists y)(x \approx_N [x \,\Delta\, y]^y)$$

$$\text{Property } (x) \ \mathit{iff}_{\mathrm{df}} \ (\exists y)(x \approx_N [z \,\Delta\, y]^y_z)$$

$$N\text{-ary Relation } (x) \ \mathit{iff}_{\mathrm{df}} \ (\exists y)(x \approx_N [\langle z_1, \ldots, z_n \rangle \,\Delta\, y]^y_{z_1 \ldots z_n})$$

$$\text{True } (x) \ \mathit{iff}_{\mathrm{df}} \ (\exists y)(x \approx_N [(\exists z) z \,\Delta\, y]^y \;\&\; (\exists z) z \,\Delta\, y)$$

$$\ldots \{x, y\} \ldots \ \mathit{iff}_{\mathrm{df}} \ (\exists z)((\forall w)(w \,\Delta\, z \equiv (w = x \lor w = y)) \;\&\; \ldots z \ldots)$$

where z is a new variable not occurring in ...

$$\langle v_1 \rangle =_{\mathrm{df}} v_1$$

$$\langle v_1, v_2 \rangle =_{\mathrm{df}} \{\{v_1\}, \{v_1, v_2\}\}$$

$$\langle v_1, \ldots, v_{n+1} \rangle =_{\mathrm{df}} \langle \langle v_1, \ldots, v_n \rangle, v_{n+1} \rangle.$$

The definition of individual particular has interesting historical roots as far back as works by Peter Abelard and Leibniz and, more recently, in Lesniewski's 'Ontology' and Quine's 'New Foundations' and *Mathematical Logic*. The definition says that x is an individual particular if and only if it is necessary that x is predicable of itself and itself only. What is particular about particulars is that necessarily they are predicable of themselves and themselves only. The definition of truth says that x is true if and only if there is a property y such that x is necessarily equivalent to the proposition that y has an instance and y does in fact have an instance. So, for example, let y be the property of being something z such that x is true (i.e., $[x \text{ is true}]^x_z$). Then, x is a proposition if and only if x is necessarily equivalent to the proposition that y has an instance. And, in turn, x is true if and only if y in fact has an instance. Incidentally, although this definition of truth is logically adequate, it is not the official definition that I will offer in §45. It is given here for illustrative purposes. Finally, if our background theory should be T1 instead of T2′, then $=$ should replace $\approx_N$ in the above definitions.

Now in order to contextually define any given second-order sentence C in the first-order language L_ω with Δ and $\approx_N$, simply convert C into the sentence C' of L_ω by means of the following conversion rules:

(1) Second-order atomic formulas (where p_i is a sentential variable and f_i^n is a predicate variable):
$p_i \Rightarrow \text{True}\,(p_i)$.

$f_i^n(t_1, \ldots, t_n) \Rightarrow \langle t_1, \ldots, t_n \rangle \, \Delta f_i^n$.

(2) Restricted quantifiers (where v_j is a new variable not occurring in $A(a_i)$):

$(\forall a_i)A(a_i) \Rightarrow (\forall v_j)(v_j$ is an individual particular $\supset A(v_j))$.

$(\forall p_i)A(p_i) \Rightarrow (\forall v_j)(v_j$ is a proposition $\supset A(v_j))$.

$(\forall f_i^1)A(f_i^1) \Rightarrow (\forall v_j)(v_j$ is a property $\supset A(v_j))$.

$(\forall f_i^n)A(f_i^n) \Rightarrow (\forall v_j)(v_j$ is an n-ary relation $\supset A(v_j))$.

To apply these conversion rules to a given higher-order sentence C, begin with the innermost formula in C and apply rules (1) and (2) in that order; then, working outward in C, repeat this process until no higher-order notation remains. The result is the sentence C' of L_ω with Δ and $\approx_N$. C is then contextually defined as follows: C *iff*$_{\text{df}}$ C'.

Since L_ω with Δ and $\approx_N$ has a single sort of variable that ranges over everything, it is actually more expressive (and in this sense, more general) than the typical higher-order language. Yet it is possible, though quite uncommon, for a higher-order language to have just one sort of variable. Therefore, greater expressive power cannot be used as a fail-safe criterion for distinguishing the first-order approach to logic from the competing higher-order approach. For such a criterion we must look to the subject/predicate distinction.

23. The Subject/Predicate Distinction

The first-order approach adopts the traditional linguistic distinction between subject and predicate, between noun and verb; the higher-order approach does not. That is, on the first-order approach an absolute distinction is made between linguistic subjects and linguistic predicates such that a linguistic subject (noun) cannot except in cases of equivocation be used as a linguistic predicate (verb) and conversely. The higher-order approach does not impose such a restriction.[5]

The distinction between linguistic subject and linguistic predicate is evident in the surface syntax of natural language. To see the distinction there, notice that English predicates, e.g., the verbs 'repeats' and 'cycles', can never (without equivocation) occur as subjects: e.g., 'repeats = cycles' is just not a sentence. Likewise, subjects, like the noun phrases '1/3' and '.333...', can never (without equivocation) occur as predicates: e.g., '1/3 .333...' is not a sentence either. By contrast, subjects and predicates, when combined with each other in the proper order, do form sentences: e.g., '.333... repeats' and '1/3 cycles' are sentences. Thus, at least as far as the surface syntax of English is concerned, there does seem to be a sharp distinction between linguistic subjects and linguistic predicates. And this is the distinction that is built into the syntax of first-order languages. In the syntax of higher-order languages, however, this distinction is glossed over.

Although in natural language predicates cannot be used as subjects, it is possible to transform predicates into legitimate subjects by means of certain abstraction operations. (The resulting linguistic subjects are complex abstract noun phrases.) So by nominalizing the verb 'repeats', we may transform it into the gerund 'repeating', and by nominalizing the verb 'cycles', we may transform it into the gerund 'cycling'. Since these nominalized expressions are legitimate linguistic subjects, they can be combined with verbs (e.g., '=' and 'is') to form sentences. Hence, e.g., 'repeating = cycling', '.333... is repeating' and '1/3 is cycling' are sentences. Nominalizations are naturally represented in first-order language by means of the bracket notation. These three sentences may thus be represented by '$[Rx]_x = [Cx]_x$', '$.333 \ldots \Delta [Rx]_x$', and '$1/3 \ \Delta \ [Cx]_x$', respectively. By contrast, in a higher-order language, where the distinction between a predicate and its nominalization is glossed over, the above three English sentences would typically be represented by '$R = C$', '$R(.333\ldots)$', and '$C(1/3)$', respectively.

What function does the subject/predicate distinction have? First, in speech the distinction shows up as follows. A subject expression is the kind of expression that functions to identify a thing about which something is to be said. A predicate expression, by contrast, functions to say something about things so identified. As Strawson might put it, subjects fix the subject matter, and predicates (verbs) do the saying. Secondly, the subject/predicate distinction plays a role in syntax. For example, in the syntax for first-order extensional language there are three primitive syntactic categories—subject, predicate, operator—and one defined syntactic category—sentence (open or closed). The definition of sentence is roughly this: subjects combine with predicates to form sentences, and operators combine with sentences to form sentences. Hence, a very natural syntax.[6] Thirdly, the subject/predicate distinction plays a role in the construction of a natural, economical semantics that tallies with the intuitive concept of meaning. Let me explain.

Consider the kind of semantics that I call *Russellian semantics*. In this semantics, unlike a Fregean semantics, there is just one fundamental kind of meaning, and the familiar semantic relations of naming and expressing are defined in terms of it, together with the syntactic notions of subject and predicate. Naming is just the restriction of the meaning relation to syntactically simple linguistic subjects:

x names y *iff*$_{\text{df}}$ x is a syntactically simple linguistic subject and x means y.

And expressing is the restriction of the meaning relation to linguistic predicates and syntactically complex expressions:

x expresses y *iff*$_{\text{df}}$ x is a linguistic predicate or a syntactically complex expression and x means y.

In a first-order language, since no linguistic predicate or formula is a linguistic subject, linguistic predicates and formulas do not, according to a Russellian semantics, name at all. This result tallies with the intuitive notion of naming. For according to the intuitive notion, predicates and sentences do not name. (What do 'runs', 'equals', 'is', 'x runs', 'Everything equals something', etc. name? Intuitively, they name nothing at all.) By contrast, in higher-order languages all predicates and sentences are also linguistic subjects. Thus, by suppressing the distinction between linguistic predicates (and sentences) and linguistic subjects, the higher-order approach yields the counterintuitive consequence that all linguistic predicates (and sentences) name something.

A related difficulty arises in connection with Frege's question of how a true sentence '$a = b$' can differ in meaning from '$a = a$'. Frege's two-kinds-of-meaning semantics is expressly designed to answer this question. In §38, however, I show that for an idealized representation of natural language Russellian semantics is every bit as adequate as Fregean semantics. The argument makes use of the fact that strings such as '$F = G$' and '$F = F$' are ill-formed in a first-order language (since linguistic predicates are not counted as linguistic subjects). But such strings are well-formed in higher-order language (since linguistic predicates are there counted as linguistic subjects). Thus, in a higher-order setting, unlike a first-order one, we need special assurances that strings such as '$F = G$' and '$F = F$' do not constitute problematic new instances of Frege's puzzle. (This is Church's worry about Russellian semantics; see §38.) In this way, our simple and natural Russellian semantics becomes problematic when we move to a higher-order setting from a first-order one. This then is one more way in which the traditional subject/predicate distinction, as it is incorporated in first-order language, plays a role in linguistic theory.

The distinction between linguistic subjects and linguistic predicates is, of course, reminiscent of Frege's distinction between object-names and function-names. There are important differences, however. One of these differences is ontological in character. According to Frege's theory, object-names name things called *objects*, and function-names name things called *functions*. Objects are what Frege calls complete (or saturated); functions are what he calls incomplete (or unsaturated). (He further distinguishes ordinary functions from functions whose values are truth values. The latter he calls *concepts*. I will suppress this distinction in the present remarks.) However, in the framework of the first-order theory of PRPs there is a far more natural ontological distinction that does much the same job as Frege's function/object distinction. What I have in mind is the distinction between things that are ontological predicates and things that are not. Something is an *ontological predicate* if and only if in principle it could be expressed by a linguistic predicate. Now let us agree that an *object* is anything that could be named by a linguistic subject. While it is true that any ontological predicate is ontologically distinctive (for it is either a property or relation), it is also true that each ontological predicate is an object. Indeed, any property or relation can simply be assigned as the value of a first-order variable.[7] Herein lies the difference between ontological predicates and Frege's functions, for on Frege's theory no function can ever be an object. And so Frege must say that the concept horse is not a concept!

How did Frege arrive at this bizarre distinction? My suspicion is that the distinction had its origin in none other than Frege's proclivity to treat the logical syntax of natural language as higher-order and, specifically, in his proclivity to treat all constants in natural language as names, including even those constants that were traditionally identified as linguistic predicates. Let me explain.

Frege was well aware of natural language phenomena such as the following: for all linguistic subjects *b* and all linguistic predicates *F* if $\ulcorner \ldots b \ldots \urcorner$ has a truth value (or makes sense), then barring equivocation $\ulcorner \ldots F \ldots \urcorner$ does not have a truth value (or sense). For example, 'Cycling is a property' has a truth value (makes sense), but 'Cycles is a property' does not. (See p. 50, Frege, 'On Concept and Object'.) When Frege sought to explain such linguistic phenomena, he arrived at an ontologically based semantical explanation. 'Cycling is a property' has a truth value (sense) because 'cycling' and 'is a property' name (express) things that by their nature combine together to yield something else; that thing is the nominatum (sense) of 'Cycling is a property'. By contrast, 'Cycles is a property' does not have a truth value (sense) because 'cycles' and 'is a property' name (express) things that cannot by their nature combine together to yield something. Hence, there is nothing with which to identify the nominatum (sense) of 'Cycles is a property'.

In contrast to Frege's ontologically based higher-order semantical explanation, the first-order explanation of the above natural language phenomena is *syntactic*. Complex expressions have truth value (make sense) if and only if they are syntactically well-formed formulas. To be a syntactically well-formed formula a complex expression must be built up according to the syntactic formation rules. However, the syntactic formation rules prohibit using linguistic predicates as linguistic subjects. This simple syntactic line of explanation was unavailable to Frege, for on his theory both linguistic subjects and predicates are names. Thus, Frege could explain the failure of the substitutability of linguistic predicates for linguistic subjects only by positing a bizarre ontological distinction between the kind of things named by linguistic predicates and the kind of things named by linguistic subjects, i.e., by positing the distinction between functions and objects.

24. The Property/Fnnction Distinction

Frege's bizarre ontological distinction between functions and objects has not had much impact historically. Nevertheless, the Fregean doctrine that predicates name functions has had a persistent influence on subsequent higher-order formulations of logic. Here, of course, the theory that functions cannot be objects is suppressed. The practice of treating predicates as naming functions has been taken up by Russell (in *Principia Mathematica*), Church, Henkin, Montague, Kaplan, and David Lewis, to name a few. I will now make some criticisms of this practice.

Consider the following intuitively valid argument:

x is red and red differs from blue.

∴ There is something that x is and it differs from blue.

The standard higher-order representation of this argument is:

$$\frac{R(x)\ \&\ R \neq B}{\therefore\ (\exists f)(f(x)\ \&\ f \neq B)}$$

where the predicates R and B are construed as names of the properties red and blue, respectively, and f is a 1-place predicate variable. Now if in accordance with the common higher-order practice (n-ary) predicates are also construed as naming (n-ary) functions, then the properties red and blue must be identified with 1-ary functions.[8] Indeed, all properties (i.e., all 1-ary intensional entities) must on this higher-order approach be identified with 1-ary functions. And similarly, n-ary relations (i.e., n-ary intensional entities, for $n \geq 2$) must be identified with n-ary functions.

But how unnatural such identifications are. Joy, the shape of my hand, the aroma of coffee—these are not functions. When I feel joy, see the shape of my hand, or smell the aroma of coffee, it is not a function that I feel, see, or smell. (For more on sensing and feeling, see §49.) Indeed, from the intuitive point of view n-ary functions are just a special kind of $n + 1$-ary relations, namely, those $n + 1$-ary relations that are univocal. Thus, the higher-order practice results in an identification of properties with 2-ary relations, 2-ary relations with 3-ary relations, 3-ary relations with 4-ary relations, etc. This outcome is entirely unintuitive.

On the first-order approach, this unintuitive outcome is easy to avoid. The above argument, for example, is straightforwardly represented as

$$\frac{x \,\Delta\, r \,\&\, r \neq b}{\therefore\; (\exists w)(x \,\Delta\, w \,\&\, w \neq b)}$$

where r and b are singular terms denoting the properties red and blue, respectively. On this approach properties are just what they should be—1-ary intensional entities. Likewise, n-ary relations-in-intension are just what they should be—n-ary intensional entities. And propositions are just what they should be—0-ary intensional entities.

The higher-order practice of identifying properties with functions has often led to another difficulty. To dramatize this difficulty consider the following propositions, where x is some particular:

$[Fx]^x$

$[x \,\Delta\, [Fy]_y]^x$

$[x \,\Delta\, [u \,\Delta\, [Fy]_y]_u]^x$

$[[Fy]_y \,\Delta\, [x \,\Delta\, v]_v^x]^x$

$[\langle x, [Fy]_y\rangle \,\Delta\, [u \,\Delta\, v]_{uv}]^x$

$[\langle [Fy]_y, x\rangle \,\Delta\, [u \,\Delta\, v]_{vu}]^x$ ....

Although on conception 1 these propositions are identical, on conception 2, which concerns intentional matters, these propositions are all distinct. Now consider any higher-order functional approach to intensional logic that does not avail itself of a primitive Δ-predicate. (If a theory does avail itself of a Δ-predicate, one can hardly see the point of making the theory higher-order; recall §22.) On such a higher-order functional approach, the above propositions would be represented:

$$F_{o1\iota}(x_\iota)$$

$$(\lambda y_\iota)(F_{o1\iota}(y_\iota))(x_\iota)$$

$$(\lambda u_\iota)((\lambda y_\iota)(F_{o1}(y_\iota))(u_\iota))(x_\iota)$$

$$(\lambda f_{o1\iota})(f_{o1\iota}x_\iota)(\lambda y_\iota)(F_{o1\iota}(y_\iota))$$

$$(\lambda f_{o1\iota})(\lambda u_\iota)(f_{o1\iota}(u_\iota))((\lambda y_\iota)(F_{o1\iota}(y_\iota)), x_\iota)$$

$$(\lambda u_\iota)(\lambda f_{o1\iota})(f_{o1\iota}(u_\iota))(x_\iota, (\lambda y_\iota)(F_{o1\iota}(y_\iota))) \quad \ldots.$$

However, given the usual laws for λ, if the above propositions are represented in this way, they would all have to be identical. Therefore, intensional distinctions relevant to the logic for intentional matters are lost on the above kind of higher-order functional approach. Where does this approach go wrong?

Without attempting a detailed analysis, I think that I can in a rough way indicate the source of the problem. Consider the first two propositions $[Fx]^x$ and $[x \Delta [Fy]_y]^x$. Given the algebraic methods developed in chapter 2, we have the following:

$$[Fx]^x = \mathrm{Pred}_0([Fy]_y, x)$$
$$[x \Delta [Fy]_y]^x = \mathrm{Pred}_0(\mathrm{Pred}_0([u \Delta v]_{uv}, [Fy]_y), x).$$

Thus, whereas the proposition $[Fx]^x$ is obtained by applying the predication operation to the property $[Fy]_y$ and x, the proposition $[x \Delta [Fy]_y]^x$ involves not only the predication operation but also the predication relation (the Δ-relation). The error in the above sort of higher-order functional approach is something like this. It in effect collapses the predication operation and the predication relation into the single Fregean operation of application of function to argument.

In view of the difficulties facing the functional approach to higher-order logic, why do higher-order theorists persist in treating predicates as names of functions rather than as names of properties? Beyond mere tradition and preoccupations with mathematics rather than natural logic, the major impetus for this practice is that it makes possible a relatively simple kind of semantics for higher-order language. A property-theoretic semantics, which would be more natural than a function-theoretic semantics, has to my knowledge never been accomplished for higher-order language. My conjecture is that the simplest way to construct one is, ironically, to translate the higher-order language into a first-order language, perhaps along the lines of §22, and then to do the property-theoretic semantics for the first-order language, perhaps along the lines of §§13–14.

25. The Origin of Incompleteness in Logic

We now return to the issue of incompleteness in logic, the issue with which this chapter began. Gödel showed that first-order number theory is incomplete. Since first-order number theory can be modeled within first-order set theory, first-order set theory is incomplete as well. A thesis of the next chapter is that first-order set theory can in turn be modeled within the first-order logic for the predication relation.[9] It follows that this logic is incomplete. Thus, in view of the results of chapter 2 we obtain the following fuller picture of the stages of completeness and incompleteness in first-order theories.[10]

COMPLETENESS AND INCOMPLETENESS IN FIRST-ORDER THEORIES

(1)	first-order quantifier logic with identity and the numerals	Complete
(2)	first-order quantifier logic with identity and extensional abstraction	
(3)	first-order quantifier logic with identity and intensional abstraction	
(4)	first-order quantifier logic with identity, the numerals, addition, and multiplication	Incomplete
(5)	first-order quantifier logic with identity and set membership and with or without extensional abstraction	
(6)	first-order quantifier logic with identity and predication and with or without intensional abstraction	

What is the origin of the incompleteness in logic? In view of (1), (2), and (3) in the above picture, the ontology of abstract entities clearly is not responsible. So in view of (4) in the above picture, one might be inclined to the view that the standard number-theoretic operations are responsible. However, if this answer is not elaborated, it is unconvincing. For on the face of it, operations from number theory do not even belong to logic *per se*, i.e., to the science of valid thinking. Similarly, in view of (5) in the above picture, one might be inclined to identify the relation of set membership as the source. However, as with operations from number theory, the relation of set membership does not on the face of it belong to the domain of logic *per se*.[11]

A thesis of chapter 6 is that all the usual operations from number theory are definable in L_ω in terms of the predication relation. And a thesis of chapter 5 is that, insofar as set theory has any utility in mathematics or empirical science, an ϵ-relation having all the properties attributed to ϵ in axiomatic set theory is definable in terms of the predication relation. Therefore, if these theses are correct and if the predication relation indeed falls within the domain of logic *per se*, then the incompleteness in logic can in this sense be traced to defined number-theoretic operations or to a defined ϵ-relation. However, since the logical character of these defined notions derives from their definability in terms of the predication relation, this relation, if it indeed belongs to the domain of logic *per se*, must be identified as the ultimate source of the incompleteness in logic. (See (6) in the picture opposite.)

Does the predication relation belong to the domain of logic *per se*? That is, is the theory for the predication relation truly part of the science of valid thinking? The answer to this question is obvious: if any theory at all ever qualifies as part of logic, the theory for the predication relation does; the predication relation is the very paradigm of a purely logical relation. This point, which has been neglected by virtually all twentieth century philosophers of logic,[12] cannot be stressed enough. The copula is a logical constant *par excellence*, and the theory for the copula is part of logic.

Therefore, my conclusion is this. It is not the infinite abstract ontology of logic, i.e., the infinite ontology of properties, relations, and propositions, that is responsible for the incompleteness in first-order logic. Rather, the ultimate source of the incompleteness is a fundamental logical relation on that abstract ontology, the predication relation. The logic for properties, relations, and propositions, the logic for L_ω is provably complete as long as no predicate is singled out as a distinguished logical predicate expressing the predication relation. However, as soon as a predicate is singled out in this way, the resulting logic is rendered incomplete.

What is the source of the incompleteness in higher-order theories? In higher-order settings, unlike the first-order setting, philosophically relevant stages of completeness and incompleteness evidently cannot be isolated,[13] for higher-order quantification theory is incomplete from the very start. Given the hypothesis that the predication relation is the source of the incompleteness in logic, we can explain the inability to separate philosophically relevant stages of completeness and incompleteness in higher-order quantification theory. This theory is incomplete from the start because the notation for the predication relation is built into the syntactic structure of higher-order languages[14] and, thus, the semantic import of this notation is never permitted to vary from one standard model to another. However, if higher-order quantification theory is treated as a derived theory constructed within the first-order logic for the predication relation (as in §22), then the source of the incompleteness in higher-order quantification theory—namely, the predication relation—becomes transparent.

26. The Logical, Semantical, and Intentional Paradoxes

The source of incompleteness in first-order logic, I have argued, is traceable to the predication relation. It should be no surprise, then, that I also hold that the predication relation lies at the heart of the familiar paradoxes that have plagued logicians over the years, e.g., the paradoxes of Russell, Cantor, Burali-Forti and the paradoxes of Epimenides, Berry, Grelling, and Richard. Specifically, I hold that, when properly analysed, each of these paradoxes involves some kind of self-refuting predication.

How do the paradoxes arise? The algebraic semantic technique provides a new perspective on this question. Consider the standard model structure $\mathscr{M}$ for L_ω with Δ:

$$\langle \mathscr{D}, \mathscr{P}, \mathscr{K}, \mathscr{G}, \mathrm{Id}, \bar{\Delta}, \mathrm{Conj}, \mathrm{Neg}, \mathrm{Exist}, \mathrm{Exp}, \mathrm{Inv}, \\ \mathrm{Conv}, \mathrm{Ref}, \mathrm{Pred}_0, \mathrm{Pred}_1, \mathrm{Pred}_2, \ldots \rangle$$

Here $\bar{\Delta}$ is the relation-in-intension in $\mathscr{D}_2$ that is expressed by the predicate Δ on its standard interpretation. Now, what would one think is the extension of the predication relation $\bar{\Delta}$? Intuitively, one would think that a pair x, y is in the extension of the predication relation $\bar{\Delta}$ if and only if x is in the extension of y. That is, one would think that $\mathscr{G}(\bar{\Delta}) = \{xy \in \mathscr{D}: x \in \mathscr{G}(y)\}$ and, more generally, that $(\forall H \in \mathscr{K})H(\bar{\Delta}) = \{xy \in \mathscr{D}: x \in H(y)\}$. But this is impossible, as the following model-theoretic analogue of Russell's paradox shows. Suppose that $\mathscr{G}(\bar{\Delta}) = \{xy \in \mathscr{D}: x \in \mathscr{G}(y)\}$.

(1)	$\mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})) \in \mathscr{G}(\mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})))$	Premise
(2)	$\mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})) \notin \mathscr{G}(\mathrm{Ref}(\bar{\Delta}))$	By (1) & Neg-rule[15]
(3)	$\langle \mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})), \mathrm{Neg}(\mathrm{Ref}(\bar{\Delta}))\rangle \notin \mathscr{G}(\bar{\Delta})$	By (2) & Ref-rule
(4)	$\langle \mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})), \mathrm{Neg}(\mathrm{Ref}(\bar{\Delta}))\rangle \notin \{xy \in \mathscr{D}: x \in \mathscr{G}(y)\}$	By (3) & hypothesis
(5)	$\mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})) \notin \mathscr{G}(\mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})))$	By (4) & set theory
(1′)	$\mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})) \notin \mathscr{G}(\mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})))$	Premise
(2′)	$\mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})) \in \mathscr{G}(\mathrm{Ref}(\bar{\Delta}))$	By (1′) & Neg-rule
(3′)	$\langle \mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})), \mathrm{Neg}(\mathrm{Ref}(\bar{\Delta}))\rangle \in \mathscr{G}(\bar{\Delta})$	By (2′) & Ref-rule
(4′)	$\langle \mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})), \mathrm{Neg}(\mathrm{Ref}(\bar{\Delta}))\rangle \in \{xy \in \mathscr{D}: x \in \mathscr{G}(y)\}$	By (3′) & hypothesis
(5′)	$\mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})) \in \mathscr{G}(\mathrm{Neg}(\mathrm{Ref}(\bar{\Delta})))$	By (4′) & set theory

Thus, given the law of the exluded middle, the hypothesis that $\mathscr{G}(\bar{\Delta}) = \{xy \in \mathscr{D}: x \in \mathscr{G}(y)\}$ leads to a contradiction.

Another way to see the difficulty is this. Given the algebraic semantics for L_ω, the following holds for all formulas A:

$$\langle \mathscr{A}(v_1), \ldots, \mathscr{A}(v_j)\rangle \in \mathscr{G}(\mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}([A]_{v_1\ldots v_j})) \textit{ iff } \mathrm{T}_{\mathscr{I}\mathscr{A}\mathscr{M}}(A).$$

Therefore, if $\mathscr{G}(\bar{\Delta}) = \{xy \in \mathscr{D}: x \in \mathscr{G}(y)\}$, then $v\,\Delta\,[A]_v \equiv A$ would have to be true for all formulas A. But this is just the principle of predication from which the property-theoretic analogue of Russell's paradox follows immediately:

$$[v \not\Delta v]_v \,\Delta\, [v \not\Delta v]_v \equiv [v \not\Delta v]_v \not\Delta [v \not\Delta v]_v.$$

What is going on? The language L_ω is *semantically complete* in the sense that, for every formula A, there is a singular term (namely, the normalized intensional abstract $[A]_\alpha$) that denotes the meaning of A. That is, all *expressible* properties, relations, and propositions are *denotable*. To my mind any language that provides an ideal treatment of modal and intentional matters ought to be semantically complete in this sense. Now consider a semantically complete language (e.g., L_ω) whose sentential and quantificational logic is classical. If such a language has a predicate (e.g., Δ) that expresses the predication relation, then necessarily the extension of the predication relation is different from what one would naively take it to be. If classical logic is sound, then, paradoxes in a semantically complete language originate in a mistake concerning the extension of the predication relation.

If classical logic is not to be tampered with,[16] then a resolution of the paradoxes in semantically complete languages must involve modifications in what one naively takes to be the extension of the predication relation. So it is quite pleasing to see that this is precisely what happens when the standard resolutions of the paradoxes in naive first-order set theory are adapted to first-order intensional logic with predication. Until we find an ideal resolution of the paradoxes of predication, we may therefore follow this maxim: to obtain a workable resolution of these paradoxes, determine the best resolution of the paradoxes in first-order set theory and then adapt it to the setting of intensional logic with predication.

For illustrative purposes I will now sketch how such adaptation works in the case of the two most familiar resolutions of the first-order set-theoretical paradoxes, namely, Zermelo's resolution and von Neumann's resolution.[17] In connection with the von Neumann-style resolution I will say that an object is *safe* if and only if it has properties, i.e.,

$$S(v_i) \text{ } \mathit{iff}_{\mathrm{df}} \text{ } (\exists v_j)\, v_i \,\Delta\, v_j.$$

As a notational convention, let the letters $a, b, c, \ldots$ be introduced as special restricted variables that range over safe things. Accordingly, $(\forall a_i)A(a_i)$ is short for $(\forall v_j)(S(v_j) \supset A(v_j))$, and $(\exists a_i)A(a_i)$ is short for $(\exists v_j)(S(v_j) \,\&\, A(v_j))$ where v_j is a new distinct variable. Now, as I have said, the following is the naive principle of predication that is responsible for the paradoxes:

(Naive Principle of Predication)
For any formula A,
$\vdash \langle v_1, \ldots, v_j \rangle \,\Delta\, [A]_{v_1 \ldots v_j} \equiv A.$

According to the Zermelo-style and the von Neumann-style resolutions of the logical paradoxes, the naive principle of predication is modified as follows:

(Zermelo-Style Principle of Predication)
For any formula A having the form $(v_1, \ldots, v_j \Delta u \,\&\, B)$,
$\vdash \langle v_1, \ldots, v_j \rangle \Delta [A]_{v_1 \ldots v_j} \equiv A$.

(von Neumann-Style Principle of Predication)
For any formula A where, for all h, $1 \leq h \leq j$, a_h is free for v_h in A and conversely,
$\vdash \langle a_1, \ldots, a_j \rangle \Delta [A(v_1, \ldots, v_j)]_{v_1 \ldots v_j} \equiv A(a_1, \ldots, a_j)$.[18]

The L_ω counterparts of the remaining Zermelo-Fraenkel (ZF) and von Neumann-Gödel-Bernays (GB) axioms—minus extensionality—are formulated on analogy.[19] By adding the ZF-style axioms or the GB-style axioms to T1 or T2′, we obtain the rudiments of four logics for L_ω with Δ.

Now what about the logical paradoxes? Evidently, the closest we can come to, e.g., Russell's paradox in the two ZF-style logics for L_ω with Δ is

$$\begin{aligned}&[x \Delta u \,\&\, x \not\Delta x]^u_x \Delta [x \Delta u \,\&\, x \not\Delta x]^u_x \\ &\quad \equiv ([x \Delta u \,\&\, x \not\Delta x]^u_x \Delta u \\ &\qquad \& \, [x \Delta u \,\&\, x \not\Delta x]^u_x \not\Delta [x \Delta u \,\&\, x \not\Delta x]^u_x)\end{aligned}$$

from which it follows merely that

$$(\forall u)([x \Delta u \,\&\, x \not\Delta x]^u_x \not\Delta u).$$

And the closest we can come to Russell's paradox in the two GB-style logics for L_ω with Δ is

$$S([x \not\Delta x]_x) \supset ([x \not\Delta x]_x \Delta [x \not\Delta x]_x \equiv [x \not\Delta x]_x \not\Delta [x \not\Delta x]_x)$$

from which it follows merely that

$$\neg S([x \not\Delta x]_x).$$

Thus, we may tentatively conclude that the above logics for L_ω with Δ are free of contradiction.[20]

But what about the semantical and intentional paradoxes? To set the stage for the discussion of these paradoxes, note the following surprising fact. In each of the above logics for L_ω with Δ it is possible to define a *truth predicate T for propositions* such that the following condition of adequacy is provable for all formulas A:

$$T[A] \equiv A.$$[21]

In view of Tarski's theorem on the undefinability within a given language of a truth predicate for the sentences of that language, the definability within the logic for propositions of a truth predicate for propositions might appear paradoxical. But it is not. To get a semantical paradox something more is required, specifically, a special interpretation of L_ω. Suppose that L_ω is interpreted in such a way that one of its primitive predicates (let it be M^2) expresses the *meaning relation* for L_ω. In that case a truth predicate Tr for the sentences in L_ω could be defined in L_ω as follows:

$$Tr(x)\ \mathit{iff}_{\mathrm{df}}\ T((\iota y)M^2(x, y))$$

i.e.,

x is a true sentence *iff*$_{\mathrm{df}}$ what x expresses is true.

Given this definition and the above condition of adequacy for T, the following condition of adequacy for Tr would hold for all sentences A in L_ω:

$$Tr\ulcorner A\urcorner \equiv A.$$[22]

And this does contradict Tarski's theorem. Therefore, if L_ω can be interpreted in such a way that one of its predicates expresses the meaning relation for L_ω, the ZF-style and GB-style principles of predication must be modified further.

In a similar vein, although the above logics for L_ω with Δ are as they stand free of intentional paradoxes, intentional paradoxes can easily be manufactured by suitably interpreting L_ω and by adjoining certain empirically conceivable auxiliary premises. For example, let L_ω be interpreted so that one of its predicates expresses an intentional relation, e.g., belief. And suppose that there is someone who believes that he is sometimes mistaken but (with the possible exception of some of his beliefs that are entailed by this one together with his true beliefs) all his other beliefs are true.[23] From this supposition it is possible to derive the following logical falsehood in the above ZF and GB-style logics for L_ω with Δ:

$$xB[(\exists y)(xBy \,\&\, \neg Ty)]^x \equiv \neg xB[(\exists y)(xBy \,\&\, \neg \mathrm{T}y)]^x.$$

Hence, an intentional paradox.

Despite the ease with which semantical and intentional paradoxes seem to be generated, such paradoxes may be neatly resolved simply by further adjusting the extension of the predication relation. Take any formula A. Consider any quantified occurrence of a variable v_i in A. Suppose that this occurrence of v_i is bound by an occurrence in A of a quantifier $(\exists v_i)$ or $(\forall v_i)$ that itself is not a constituent of an occurrence in A of (the expanded form of) our definition of the truth predicate T.[24] Such occurrences of variables in A will be called *ungrounded*. Let A_u be the formula that results from restricting the range of ungrounded occurrences in A to things that have u as a property, and let such formulas A_u be called *grounded*. Now consider the following modified principles of predication:

(ZF-Style Predicative Principle of Predication)
If A_u has the form $(v_1, \ldots, v_j \Delta u \,\&\, B)$, then
$\vdash \langle v_1, \ldots, v_j \rangle \Delta [A_u]_{v_1 \ldots v_j} \equiv A_u$.

(ZF-Style Impredicative Principle of Predication)
If w is distinct from $v_1, \ldots, v_j$ and is not free in A and if A has the form $(v_1, \ldots, v_j \Delta u \,\&\, C)$, then
$\vdash (\exists w)(\langle v_1, \ldots, v_j \rangle \Delta w \equiv_{v_1 \ldots v_j} A)$.[25]

(GB-Style Predicative Principle of Predication)
If for all h, $1 \leq h \leq j$, a_h is free for v_h in A_u and conversely, then
$\vdash \langle a_1, \ldots, a_j \rangle \Delta [A_u(v_1, \ldots, v_j)]_{v_1 \ldots v_j} \equiv A(a_1, \ldots, a_j)$.[26]

(GB-Style Impredicative Principle of Predication)
If w does not occur in A, then
$\vdash (\exists w)(\langle a_1, \ldots, a_j \rangle \Delta w \equiv_{a_1 \ldots a_j} A)$.

Let the L_ω counterparts of the remaining ZF and GB axioms (minus extensionality) be formulated on analogy.[27] By adding the modified ZF-style axioms or the modified GB-style axioms to T1 or T2′, we obtain four logics for L_ω with Δ. Evidently, none of the familiar semantical or intentional paradoxes can be generated in these modified ZF-style and GB-style logics even when a univocal meaning predicate and various intentional predicates are singled out.[28] And at the same time, we still can define the univocal truth predicate T such that the following modified condition of adequacy is provable for all grounded formulas A_u:

$$T[A_u] \equiv A_u.$$

What is the intuitive idea behind this resolution of the semantical and intentional paradoxes? It is that in all contexts of speech and thought there is an implicit limitation *u* on the things that are taken to be relevant for consideration. That is, in all contexts of speech and thought an implicit universe of discourse *u* is invoked, where *u* is something less than the totality of all things. In a given context the identity of *u* is determined pragmatically by features of the context. The semantical and intentional paradoxes result from a failure to notice and keep track of subtle contextual shifts affecting the implicit universe of discourse.[29]

The idea that the semantical and intentional paradoxes can be resolved by making explicit contextually invoked limitations on the universe of discourse ought to sound familiar. For the ramified theory of types embodies a special case of this very idea. Indeed, the modified ZF-style and GB-style logics for L_ω with Δ may be viewed as natural generalizations of ramified type theory.[30] However, these logics for L_ω with Δ are generalizations that eliminate most of the artificiality and rigidity for which ramified type theory is notorious. The bearing this fact has on the first-order/higher-order controversy, with which this chapter has been concerned, is that ramified type theory is typically formulated as a higher-order logic. So once again the naturalness and generality of first-order logic comes through.

Is higher-order logic best viewed as a natural generalization of first-order logic, or is it best viewed as an artificially restricted theory derived within first-order logic with predication? The answer, I hope, is evident.

The proposed resolution of the semantical paradoxes depends essentially on the fact that intensional entities are the primary semantical correlates of formulas. No analogous resolution is possible if instead extensional entities—namely, sets—are identified as the primary semantical correlates.[31] This problem in set-theoretical semantics is just the beginning of the troubles for a formal philosophy based on set theory. In the next chapter we shall find many more.

5

Class

> The notion of class is so fundamental to thought that we cannot hope to define it in more fundamental terms.
>
> W.V.O. Quine
> *Set Theory and Its Logic*

The above passage typifies the attitude most set theorists take towards their subject. From the point of view of the theory of properties, relations, and propositions, however, this attitude has two flaws. First, the notion of class is not fundamental to thought. And secondly, insofar as the notion of class is useful in mathematics and empirical science, it can be defined in more fundamental terms, namely, in terms of the predication relation. These are striking criticisms not likely to be accepted without support. The purpose of the present chapter is to provide that support.

What justifies the ontology of sets? In chapter 1, I argued that the theory of PRPs is part of logic. Since logic always belongs to the best comprehensive theory of the world, the ontology of PRPs is justified. In view of the youthfulness of set theory, however, it would be unwise to assume that the same is true for sets. We should entertain the hypothesis that set theory is the result of conflating certain constructions that, although they do play a role in the logic of natural language, do not play the role that set theory presumes. I am inclined to this hypothesis and, indeed, to the proposition that there is simply no sound justification for the ontology of sets. How might the set theorist attempt to justify his ontology? There are three strategies open to him. The first is to show that sets are included in what might be called our *naturalistic ontology*. If they are, then we may assume that whatever justifies our naturalistic ontology also justifies the ontology of sets. The second strategy is to show that, like the theory of PRPs, set theory is part of logic. In this case, the ontology of sets would be justified in the same way as the ontology of PRPs. And the third strategy is to show that set theory plays some unique role in mathematics or in empirical science. If it does, then its ontology would be justified pragmatically. None of these strategies is successful, however, as I will now explain.

27. The Unnaturalness of Sets

Paul Halmos begins his popular book *Naive Set Theory* with this observation:

> A pack of wolves, a bunch of grapes, or a flock of pigeons are all examples of sets of things.

Perhaps it is true that the idea of a set is somehow "genetically" related to ideas of such naturalistic objects as packs, bunches, and flocks.[1] Nevertheless, it is certain that sets are not the same sort of thing as packs, bunches, flocks, etc. Here are a few of the many reasons. First, packs, bunches, flocks, tribes, and so on, displace volumes, have mass, and come into and pass out of existence. Sets, by contrast, are non-physical and eternal. Secondly, sets cannot change their members; packs, bunches, flocks, etc. can. If a wolf in* a given pack dies (or gives birth), the pack is still the same pack. But the set of wolves-before-the-death (birth) is not the same set as the set of wolves-after-the-death (birth). Thus, a set of wolves and a pack of wolves are different. Thirdly, packs, bunches, flocks, etc. do not exist if nothing is in them; this is not so for sets. If there were no wolves, there would be no packs of wolves. But the set of wolves would exist nonetheless, for it would just be the null set. Indeed, if sets exist, the null set is a set that exists necessarily.[2]

If sets are not the same sort of thing as packs, bunches, flocks, etc., what are they? It is now commonplace to say that sets are collections or classes. What is meant by this? Art collections, social classes, sets of dishes: is it true that these are cases of the kind of sets posited in set theory? No, definitely not. They are no more the kind of sets posited in set theory than are packs, bunches, and flocks, etc., and for much the same reasons. First, art collections and sets of dishes can displace volumes, have mass, and come into and pass out of existence. And social classes, although they seem not to displace volumes or have mass, can come into and pass out of existence. Secondly, art collections, social classes, sets of dishes,

* Note that in order not to bias the discussion I will use the natural and, I hope, neutral locution 'is in' (and its cognates) rather than the technical locution '$\in$'. A moment's reflection will show that in adopting this practice I do not commit any fallacies of equivocation.

etc. can change their members.[3] Thirdly, ordinary collections, social classes, and ordinary sets do not exist if nothing is in them. (If China has no aristocrats, it has no aristocracy.)

These three differences suffice to show that ordinary collections, social classes, and ordinary sets are different from set-theoretical sets. However, there might be a fourth difference, having special philosophical interest. This difference concerns a transitivity property. Consider a billionaire who collects art collections in the style in which Howard Hughes used to collect companies. This man purchases outright entire art collections. Now if his, say, ten art collections contain one Cezanne each, then we would say that there are ten Cezannes in his collection of art collections. And in general, if a painting is in an art collection that is itself in a collection of art collections, then we would say that the painting itself is in the collection of art collections. The sets of set theory are not like this at all. No individual paintings are in the set of art collections; only art collections are.[4] Thus, the set of art collections and the collection of art collections are different. This sort of difference also seems to hold between set-theoretical sets, on the one hand, and social classes and ordinary sets, on the other. For example, if Jones is in the intelligentsia and the intelligentsia is in the upper class, then we would say that Jones is in the upper class. Or if a saucer is in a matched cup-and-saucer set that is itself in a set of eight matched cup-and-saucer sets, then we would say that the saucer is in the set of cup-and-saucer sets. And we say that there are four socks in a pair of pairs of socks. None of these things hold for the set-theoretical counterparts—the set of upper classes, the set of sets containing a cup and a matching saucer, the set consisting of $\{\text{sock}_1, \text{sock}_2\}$ and $\{\text{sock}_3, \text{sock}_4\}$. Put formally, the difference here is that ordinary collections, social classes, and ordinary sets seem transitive whereas the sets of set theory typically are not. That is, the following transitivity principle seems to be valid for ordinary collections, social classes, and ordinary sets whereas it is not valid for the sets posited in set theory:

$$x \text{ is in } y \supset (\forall z)(z \text{ is in } x \supset z \text{ is in } y).$$

This transitivity principle is equivalent to the following: $(\exists x)(z$ is in x & x is in $y) \supset z$ is in y. Ordinary collections, social classes, and ordinary sets would thus seem to be closed under a union operation.

There might be yet another difference between the sets of set theory and ordinary collections, social classes, and ordinary sets, one which concerns a corresponding power operation. Consider an example. If the individual cup and individual saucer in a matched cup-and-saucer set are themselves in a full set of dishes, then we say that the matched cup-and-saucer set itself is in the set of dishes. The sets posited by set theory are not like this. The set-theoretical set of dishes contains only individual dishes, not cup-and-saucer sets. For another example, suppose that I have a collection of famous rare stamps known as the First Issue Collection. This collection contains rare stamps from a wide variety of countries. Suppose further that I have a particularly valuable collection of rare Dutch stamps. Now if every stamp in my collection of Dutch stamps is in the First Issue Collection, we would say that my Dutch stamp collection is in the First Issue Collection. But the set-theoretical set of stamps that I own contains only stamps; it does not contain, e.g., the set of Dutch stamps that I own. To put this formally, the following power principle might be valid for ordinary collections, social classes, and ordinary sets whereas it is not valid for the kind of sets posited in set theory:

$$(\forall z)(z \text{ is in } x \supset z \text{ is in } y) \supset x \text{ is in } y$$

for all x and y, where $x \neq y$.

It is commonplace among historians of logic and mathematics to remark that it was not until well into the nineteenth century that people became clear about the significant difference between membership and inclusion. However, given the above principles of transitivity and power, it follows that for ordinary collections, social classes, and ordinary sets these relations are virtually equivalent; i.e., for ordinary collections, social classes, and ordinary sets x and y:

$$x \text{ is in } y \equiv (\forall z)(z \text{ is in } x \supset z \text{ is in } y)$$

where $x \neq y$. In addition, this principle would seem to hold for at least certain ordinary collections and ordinary sets x and y, where $x = y$. In view of this, it might be more accurate to say that it was not until well into the 19th century that people became confused about the nature of membership and inclusion relations. For it was not until the set theorists' distinction was thought up that the commitment to the new, extraordinary kind of collection was made official. For that matter, it was not until the set theorists' distinction was thought up that it became possible to generate the paradoxes of naive set theory. Without the set-theoretical distinction between membership and inclusion there would be no set-theoretical paradoxes. What I mean by this will become clearer below.

By abstracting from the intuitive notions of ordinary collections, social classes, and ordinary sets as characterized in the foregoing discussion, one arrives at the general notion of what I will call an *aggregate*. Aggregates are like sets in that whenever a thing w satisfies a formula A, w is in the set of As, and w is in the aggregate of As. That is, the following schemas hold for sets and aggregates, respectively:

(a) $A(w) \supset w$ is in the set of things y such that $A(y)$
(a′) $A(w) \supset w$ is in the aggregate of things y such that $A(y)$.

Furthermore, whenever something is in the set of As it also satisfies the formula A. That is, the following converse of (a) holds for sets:

(b) w is in the set of things y such that $A(y) \supset A(w)$.

Here aggregates part company with sets, however. Recall that membership and inclusion are virtually equivalent for aggregates. Thus, if a thing w is in the aggregate of As it does not follow that w satisfies A; w may instead be in something else that satisfies A, *or* something else that satisfies A could be in w, *or* something else that is in w could be in some third thing that satisfies A—any of these alternatives would do equally well. So, the schema for aggregates that corresponds to (b) offers several alternatives:

(b′) w is in the aggregate of things y such that $A(y) \supset (A(w)$ or
$(\exists u)(w$ is in u & $A(u))$ or
$(\exists u)(u$ is in w & $A(u))$ or
$(\exists u, v)(u$ is in w & u is in v & $A(v)))$.

Membership and inclusion are quite distinct in set theory, of course. So the consequent of (b) contains none of the supplementary alternatives that we had to add in (b′); whatever is in the set of As must satisfy A—there is no alternative. This dissimilarity of schemas (b′) and (b) is, of course, one more difference between ordinary aggregates and sets. But what is more important is that this feature of naive set theory is the very feature that renders it inconsistent. Schema (b) requires that all things in the set of things satisfying a given formula (e.g., the formula 'y is not in y') must themselves satisfy the formula, and this is what plunges naive set theory into contradictions. Because (b′) does not likewise restrict the identity of the things that are in ordinary aggregates, the theory of ordinary aggregates avoids the fate of naive set theory.[5]

The theory of aggregates is from a formal point of view rather like Leśniewski's mereology (i.e., the part/whole calculus);[6] each of the principles for aggregates also holds for mereological sums. In this, ordinary collections, social classes, and ordinary sets are far closer to mereological sums than to abstract sets. Set theory just does not get its motivation from the naturalistic ontology of ordinary collections, social classes, and ordinary sets.

The moral is that sets are not in evidence in any of the above naturalistic ontologies. Those who persist in the attempt to motivate the concept of class along such naturalistic lines sooner or later find themselves offering the "invisible-plastic-bag" conception. But this, I think, only confirms the point that sets do not fall within our naturalistic ontology.

28. No Basis in Logic

The second candidate strategy for justifying the ontology of sets is to attempt to show that set theory is grounded in logic. The most promising line is to look for evidence that set theory is embedded in the logical syntax of natural language. I can think of only one syntactic construction in natural language that might fill the bill, namely, pluralization.[7] Let us see how the set theorist might try to show that set theory has a special role to play in the treatment of plurals.

Consider the following sentences:

(1) The walnuts outweigh the pecans.
(2) The counties outnumber the states.

These sentences are not transformed universal conditionals:

(1′) $(\forall x, y)((\text{Walnut}(x) \;\&\; \text{Pecan}(y)) \supset \text{Outweigh}\,(x, y))$
(2′) $(\forall x, y)((\text{County}(x) \;\&\; \text{State}(y)) \supset \text{Outnumber}\,(x, y))$.

For whereas (1′) and (2′) are false, (1) and (2) are true.

Provisionally, then, let us represent (1) and (2) as 2-place relational sentences:

(1″) Outweigh (the walnuts, the pecans)
(2″) Outnumber (the counties, the states).

Here the plurals are provisionally treated as (defined or undefined) singular terms. It becomes appropriate, then, to ask what the primary semantical correlates of these provisional singular terms are. A natural hypothesis is that in (1) the primary semantical correlates of 'the walnuts' and 'the pecans' are aggregates of the ordinary sort characterized earlier (specifically, the aggregate of all walnuts and the aggregate of all pecans). On the face of it, this hypothesis seems successful. This gives rise to the presumption that the plurals in (2) should be treated analogously; i.e., this suggests that the primary semantical correlates of 'the counties' and 'the states' in (2) are also aggregates. But what kind of aggregate? Not ordinary aggregates, certainly. Since the ordinary aggregate of the counties is identical to the ordinary aggregate of the states, (2) would be false. Yet on its primary reading (2) is true. Therefore, if one continues to be swayed by the presumption that the primary semantical correlates of 'the counties' and 'the states' are aggregates, then a new, extraordinary kind of aggregate must be hypostasized. These new, extraordinary aggregates should differ from ordinary aggregates in at least the following respect: the things in the extraordinary aggregate of *F*s must be exactly those things that satisfy the predicate *F*. But this is precisely what is required of sets according to the abstraction principle of naive set theory (recall schemas (a) and (b) in §27). This gives rise to the further presumption that the extraordinary aggregate that is the primary semantical correlate of the plural 'the *F*s' in sentences akin to (2) is a set, specifically the set of *F*s.

Although the above line of reasoning has a certain appeal, it leads immediately to a fatal dilemma. Consider the following problematical sentences:

The walnuts *both* outweigh and outnumber the pecans.

Although the counties occupy exactly the same territory as the states, *they* outnumber the states, and, in addition, *they* resent federal intervention more than the states do.

These French stamps were once in the First Issue Collection;

however, after a while *they* outnumbered the Dutch stamps and, for that reason, *they* were moved to another collection.

The whales once outnumbered the human beings; now, however, *they* are nearly extinct.

In view of the earlier discussion about the nature of set-theoretical sets, if the plurals in these problematical sentences are treated in the same kind of naive surface-syntactical way adopted above in connection with sentence (2), then their primary semantical correlates clearly cannot be sets. (For example, the set of walnuts cannot outweigh the set of pecans since no set weighs anything.) These primary semantical correlates would have to be some further kind of entity. But in this case, uniformity requires us also to identify the primary semantical correlates of the plurals in (2) not with sets but with this further kind of entity. So if the plurals in the above problematical sentences get the naive surface-syntactical treatment that we provisionally gave to (2), then what initially seemed to be a justification for set theory in the natural logic of (2) evaporates. On the other hand, suppose the plurals in the above problematical sentences are treated in a sophisticated deep-structural way.[8] In this case, we nullify the original presumption that the plurals in (2) ought to be treated on analogy with the plurals in (1) (i.e., the presumption that the plurals in (2) are singular terms whose primary semantical correlates are some sort of aggregates). This makes (2) fair game for alternate sophisticated treatments; the various treatments of (2) must compete on their own terms. But if the contest is to take place in this stark arena, then, as I show next, set theory cannot win for itself a place in natural logic. Thus, either way, set theory fails to find motivation in the treatment of plurals in sentences like (2) in natural language.

To complete the above argument I must show that, if there is no presumption in favor of a set-theoretical treatment of sentences such as (2), then the set-theoretical treatment succumbs to superior competitors. So as not to bias the argument, let us agree to represent (2) provisionally along the following lines:

(2‴) $\text{Outnumber}(\{x: Cx\}, \{x: Sx\})$.

Here $\{x: Cx\}$ and $\{x: Sx\}$ are extensional abstracts; that is, they are (defined or undefined) abstract singular terms for which the following general law holds:

(3) $\{x: Ax\} = \{x: Bx\} \equiv (\forall x)(Ax \equiv Bx)$.[9]

Further, let us allow that for all non-paradox-producing formulas Ax in which y is free for x: (4) $y \in \{x: Ax\} \equiv Ay$. And finally, let us allow that (2‴) is true if and only if there is no 1-1 function from $\{x: Sx\}$ onto $\{x: Cx\}$ though there is a 1-1 function from $\{x: Sx\}$ into $\{x: Cx\}$. In this case (2‴) comes out true, as desired. Next consider briefly what seems to me to be the intuitive picture of the semantics for natural language. According to this picture, predicates and formulas do not refer to anything; they simply express.[10] A formula A, for example, expresses the property, relation, or proposition denoted by a certain associated gerundive phrase, infinitive phrase, or 'that'-clause formed from A. Specifically, it expresses the property, relation, or proposition denoted in L_ω by the normalized singular term $[A]_\alpha$. Now for all non-paradox-producing formulas A, the following law holds: (5) $\alpha \,\Delta\, [A]_\alpha \equiv A$. In view of this, the extensional abstract $\{v_i: A\}$ can be contextually defined in terms of the predication relation:

(6) $\ldots\{v_i: A\}\ldots$ *iff*$_{\text{df}}$ $(\exists v_j)((v_i \,\Delta\, v_j \equiv_{v_i} A) \,\&\, \ldots v_j \ldots)$

where v_j is a new distinct variable.[11] And $\in$ may be contextually defined as follows: (7) $u \in v$ *iff*$_{\text{df}}$ $u \,\Delta\, v$. To be convinced of the adequacy of these contextual definitions, notice that, for all non-paradox-producing formulas A and B, the above law (3) follows directly from (5) and (6), and law (4) follows directly from (5), (6), and (7). However, these laws are all that are needed for an adequate treatment of sentences such as (2‴).[12] Thus, extensional abstracts, and sentences such as (2‴), can be adequately treated within the logic for the predication relation, a theory already part of natural logic. And, this is accomplished without having to hypostasize the extraordinary aggregates of set theory. So if there is no presumption in favour of the set-theoretical treatment of sentences such as (2), then as far as natural logic is concerned the outlined alternative treatment wins hands down.

It might be objected that no economy follows from adopting this contextual treatment of extensional abstracts since sets have already entered the picture through an independent pathway, namely, through extensional semantics.[13] According to Frege's semantical theory, all meaningful expressions have two kinds of meaning: sense and nominatum. Frege identified the nominata of predicates (and open sentences) with what he called functions. But since at least the time of Tarski's work in extensional semantics, it has been common instead to view the nominatum of a predicate (open sentence) as a set, namely, the set of things that satisfy the predicate (open sentence). That is, on this view the nominatum of the predicate F is the set of Fs. Since extensional semantics already makes use of sets here, no economy of theory is gained (so someone might argue) by giving extensional abstracts such as $\{x: Fx\}$ the alternative treatment. In fact, for those persuaded by this set-theoretical semantical theory, it is only natural to identify the primary semantical correlate of the extensional abstract $\{x: Fx\}$—and thus that of the plural 'the Fs'—with the nominatum of the predicate F, i.e., with the set of Fs.

This objection, it seems to me, has gotten the proper order of the argument turned around. What good reason is there for accepting the extensional semantical theory? After all, the natural, intuitive picture of the semantics for predicates and formulas is Russell's, not Frege's. According to this picture predicates and formulas do not name anything; they simply express. The primary semantical correlate of predicates and formulas are just the properties, relations, and propositions expressed by them. What point, then, is there in having a Fregean two-kinds-of-meaning semantics rather than the simpler, more natural Russellian one-kind-of-meaning semantics? Surely something is gained at least theoretically? No, in fact, as I will show in §38, a Fregean theory provides no more semantical information than its simpler Russellian counterpart. So one can hardly justify a set-theoretical treatment of extensional abstracts and plurals by appealing to the set-theoretical content in an unnatural and informationally superfluous semantical theory.

One wonders, then, why set theory and set-theoretical semantics have caught on. Sociology of knowledge aside, one might give a "genetic" account something like the following. We have seen that there is *prima facie* evidence that plurals behave rather like singular terms. If they are singular terms, though, what are their primary semantical correlates? Some plurals seem to have ordinary concrete aggregates as their primary semantical correlates. (Recall (1) above.) This fosters the presumption that the primary semantical correlates of all plurals are aggregates. If this were so, however, then for some uses of plurals a new kind of abstract aggregate would have to be hypostasized—or so the set-theorist reasons—a kind of aggregate whose members are exactly the things satisfying the predicate from which the plural is generated. (Recall (2) above.) Thus, unlike ordinary concrete aggregates, this new abstract aggregate must wear on its sleeve the satisfaction conditions of the generating predicate. Since the new aggregate, the set of *F*s, appears to bear such a simple and direct semantical relation to the generating predicate *F*, there is a further tendency to identify the set of *F*s as the primary semantical correlate of the predicate *F* itself, as well as of the plural 'the *F*s'. And so one might arrive at the full logico-semantical belief that the set of *F*s is the primary semantical correlate of the predicate *F*.

On this account set theory and set-theoretical semantics appear to be fostered by a compulsion to concretize; that is, they appear to spring from a compulsion to think of the primary semantical correlates of predicates on analogy with ordinary concrete aggregates. Yet, as I have said, intuitively the primary semantical correlates of predicates are simply the properties or relations they express. All other semantical correlates of predicates are derivative and, furthermore, are typically natural or social:[14] packs, bunches, flocks, tribes, races, species, kinds, etc., or ordinary collections, social classes, ordinary sets, etc. Against this background set theory is seen to be an artifice out of place in the natural logical world. How ironic that a compulsion to concretize gives birth to the most abstract artifice ever produced by the human mind.

What then is the overall conclusion so far? On the basis of the foregoing critical survey it appears that the ontology of sets does not fall within our naturalistic ontology and also that set theory is not embedded in natural logic. Thus, contrary to received opinion, the notion of class does not appear to be fundamental to thought.

There remains one more strategy by which one might try to justify the ontology of sets: perhaps set theory, while not natural, is at least uniquely useful in mathematics or in empirical science. To this pragmatic issue I turn next.

29. The Dispensability of Sets

If the concepts of set theory are grounded neither in our naturalistic ontology nor in natural logic and if concepts of set theory arise from a conflation of the concepts of ordinary aggregate and property, why should one take set theory seriously? Evidently the only remaining reason is that set theory might nevertheless play some unique role in pure mathematics or in the empirical sciences.

Consider pure mathematics first. Here set theory is used in an entirely abstract way to aid and to unify the study of such matters as cardinality, order, mapping, etc. Let x be an arbitrary non-empty set. I will say that x_0 is an *ultimate element* of x if and only if $x_0 \in x_1 \in x_2 \in \ldots \in x$ and nothing is in x_0 itself. Now, it is a matter of complete indifference what the ultimate elements are of any set that might be contemplated in pure mathematics. Hence, as far as pure mathematics is concerned, the study of sets can be limited to those sets whose only ultimate element is the null set. The theory of such sets is called *pure set theory*. We may conclude, therefore, that if set theory should turn out to have a unique role to play in pure mathematics, that role can be filled by pure set theory.

Now consider the empirical sciences. Let us suppose that in the service of classification and measurement there are occasions when it is useful to consider collectively (as well as individually) the individuals with which a given empirical science deals. Let us call the sets postulated for these purposes *empirical sets*. In connection with measurement there might, in addition, be a call for certain key relations, such as equinumerosity, that hold between pure and empirical sets. The theory that characterizes empirical sets and those key relations holding between pure and empirical sets may be called *applied set theory*. If set theory has any role to play in the empirical sciences, that role can be filled by applied set theory.

So, to repeat our earlier question, why should set theory be taken seriously? The answer stated more precisely is this: pure set theory might have a unique role to play in pure mathematics, or applied set theory might have a unique role to play in the empirical sciences.

I think that pure and applied set theory have no such unique roles. In fact, this claim can be proved. Specifically, it can be proved that first-order pure and applied set theory can be modeled within the first-order logic for the predication relation. (Since first-order pure set theory countenances sets of sets, sets of sets of sets, etc., this result goes well beyond Russell's no-class construction, which works only for sets of non-sets. In what follows I will give no-class constructions for both of the leading first-order pure set theories, Zermelo-Fraenkel (ZF) and von Neumann-Gödel-Bernays (GB).[15]) This result shows that any theoretical tasks that pure and applied set theory perform can be accomplished equally well by a theory that, unlike set theory, has a legitimate origin in natural logic.[16]

Thus, my larger conclusion is that neither naturalistic ontology, natural logic, pure mathematics, nor the empirical sciences provide any ground for believing that sets exist: there is neither naturalistic, logical, nor pragmatic warrant for set theory. Set theory does not belong in a rational view of reality.

An example will help clarify what I meant when I said that first-order pure and applied set theories can be modeled within the first-order logic for the predication relation and that this shows that any of the theoretical tasks that these set theories perform can be accomplished equally well by the logic for the predication relation. Consider the miniature theory for ordered pairs:

$$\langle u, v \rangle = \langle x, y \rangle \equiv (u = x \;\&\; v = y).$$

It is widely known that this theory can be modeled by set theory. What this means is that from a syntactical point of view the notation used in the ordered-pair theory can be introduced into the language of set theory as an abbreviation for longer set-theoretic locutions and that this can be done in such a way that the miniature theory can then be derived as a theorem using just the original axioms and rules of set theory. Thus, on those occasions in pure mathematics and empirical science when previously one spoke of ordered pairs one may now merely speak of certain sets, namely, those sets singled out by the abbreviation scheme with which the ordered-pair notation is introduced. Thus, any mathematical or scientific jobs that ordered-pair theory can do can be done by set theory equally well.

In an analogous manner, then, I will show that from a syntactic point of view the notations used in first-order pure and applied set theory can be introduced into the logic for the predication relation as abbreviations for longer property-theoretic locutions and that this can be done in such a way that the set-theoretical axioms and rules can be derived as theorems using just the logic for the predication relation. This permits one to speak merely of properties on those occasions in pure mathematics and in empirical science when one previously spoke of sets. Thus the logic for the predication relation may fill perfectly well any of set theory's mathematical or empirical scientific roles.

There are two opposing philosophical purposes one might have in modeling one theory within another. One purpose is *reduction*. In the ordered-pair case, e.g., one's aim might be to show that no mathematical or scientific utility is lost if ordered pairs are *identified* with a certain kind of set. The other purpose is *elimination*. Thus, in the ordered-pair case one's motive might be to show that no mathematical or scientific utility is lost if ordered pairs are held *not to exist*.

Which of these purposes should one have, reduction or elimination? The possibility of modeling one theory within another shows that the modeled theory has no mathematical or scientific utility not possessed by the modeling theory. Suppose that the motivation offered for particular axioms in the modeling theory is at least as strong as the motivation offered for those in the modeled theory. And suppose that we already have good philosophical or logical reasons for accepting the ontological framework of the modeling theory. In this case the decision whether to reduce or eliminate the entities of the modeled theory should be based on whether there is any independent philosophical or logical reason to think those entities exist. Now, in a moment we shall see that the motivation offered on behalf of the axioms for the predication relation is at least as strong as the motivation offered on behalf of the axioms for the ϵ-relation. Further, we have already seen that there are good philosophical (§5) and logical (§§6–9) reasons for accepting an ontology of PRPs. And we have seen (in the previous two sections) that there are no independent philosophical or logical reasons for accepting an ontology of sets. Therefore, the fact that set theory can be modeled within the logic for the predication relation supports the decision to eliminate sets from our ontology.

This brings us to the motivation for the axioms in formulations of the logic for the predication relation. The point that needs to be made here is that, for any credible motivation that can be given for a particular formulation of set theory, an analogous motivation, which is at least as satisfactory, can be given for the axioms in a corresponding formulation of the logic for the predication relation. To see how this goes for a simple example, consider the usual motivation offered in support of Zermelo's axioms for pure set theory, namely, the motivation provided by the iterative conception of set. On this conception, sets are thought of as being "formed" in stages from the null set $\emptyset$ by means of repeated applications of a power operation:

Stages	1	2	...	α	...
Pure Sets	$\{\emptyset\}$	$\{\emptyset, \{\emptyset\}\}$	...	$\{y$: y is a set and every element of y belongs to a set formed prior to $\alpha\}$	...

If for every stage there is a later stage immediately following no stage, then the union of these sets is a model for Zermelo's axioms.[17] However, on analogy with the iterative conception of set, there are also iterative conceptions of PRPs. The easiest to describe is the iterative conception of pure L-determinate type 1 properties. (x is L-determinate *iff*$_{\text{df}}$ $\Box(\forall y)(y \Delta x \supset \Box y \Delta x)$.) On this conception such properties may be thought of as being "formed" in stages from the necessarily null type 1 property Λ by means of repeated applications of a power operation:

Stages	1	2	...	α	...
Pure Properties	$[\Lambda]$	$[\Lambda, [\Lambda]]$	...	$[y$ is an L-determinate property whose instances are instances of a property formed prior to $\alpha]_y$	...[18]

If for every stage there is a later stage immediately following no stage, then the union of these properties is a model for the axioms of a Zermelo-style theory for pure L-determinate type 1 properties. Moreover, the same sort of thing can be done for other iterative conceptions of PRPs.

Now the general point is this. Whenever motivation is offered for the axioms in a given set theory, it is never stronger than an analogous motivation for the axioms in an associated property theory. For this reason, general philosophical and logical considerations (such as those given earlier in this chapter) should every time guide us to choose the property theory—with its no-class construction—over the set theory.

Before I proceed to the no-class constructions for ZF and GB a general point of clarification is in order. Many of the formal metatheoretic constructions in this book are given within a set-theoretic framework, and one might wonder whether it is consistent to conduct metatheoretic constructions within a set-theoretic framework while denying that sets really exist. It is. For each of these metatheoretic constructions may be viewed as only a convenient shorthand for a metatheoretic construction within a property-theoretic framework. That is, in the last analysis my underlying metatheoretic framework is really property theory, not set theory.

30. Pure Set Theory Without Sets

There are many attractive no-class constructions of pure first-order set theory, some better suited to one philosophical view than another. The construction I will give, however, is the simplest I can find. I begin with the following definitions:

> x ultimately comprehends y *iff*$_{\mathrm{df}}$
> $(\forall z)((x \subseteq z \;\&\; (\forall w)(w \,\Delta\, z \supset w \subseteq z)) \supset y \,\Delta\, z)$
>
> x is a pure L-determinate property *iff*$_{\mathrm{df}}$
> x is an L-determinate property & whatever x ultimately comprehends is an L-determinate property.[19]

Thus, x ultimately comprehends y if and only if y is an instance of x or y is an instance of an instance of x or y is an instance of an instance of an instance of x or .... And x is a pure L-determinate property if and only if x is an L-determinate property whose instances are L-determinate properties and whose instances have as instances only L-determinate properties and so on. Now consider any sentence A in the standard language of pure first-order set theory,[20] and let A' be the sentence that arises from A by replacing all occurrences of $\in$ with Δ and by relativizing all quantifiers to pure L-determinate properties. Then I contextually define A in L_ω with Δ as follows: A *iff*$_{\mathrm{df}}$ A'. Next take the standard axioms of Zermelo-Fraenkel set theory; drop the axiom of extensionality, and rewrite the remaining axioms using Δ, $=$, and intensional abstraction.[21] TZF^- is the intensional logic obtained when these axioms are adjoined to T1.[22]

> *Metatheorem:* Every sentence that is a theorem of Zermelo-Fraenkel set theory is, given its contextual definition in terms of Δ, $=$, and intensional abstraction, a theorem of the intensional logic TZF^- (i.e., for every set-theoretical sentence A, if $\vdash_{ZF} A$, then $\vdash_{TZF^-} A$).[23]

Proof. First, we prove in TZF^- that each property is included ($\subseteq$)

in some Δ-transitive property. Take the union of the original property, the union of the result, the union of that union,.... The union of all these unions is a Δ-transitive property that includes the original property. To prove that this new property exists, we form an appropriate function on an intensional ω by means of a recursion principle (whose proof in TZF^- does not require extensionality). This function's range, which we obtain by means of the replacement axiom, is the desired Δ-transitive property. Using this theorem, we can then prove in TZF^- that each L-determinate property whose instances are all pure L-determinate properties is itself a pure L-determinate property. After this, we show by induction that $\vdash_{\mathrm{TZF}^-} A \equiv \Box A$ for formulas A whose constituent predicates are Δ or $=$ (or both Δ and $=$) and all of whose constituent terms are variables whose ranges are restricted to pure L-determinate properties. With these facts at hand, the derivation of each ZF set-existence axiom is straightforward; we simply use the associated TZF^- property-existence axiom plus appropriate instances of the TZF^- comprehension schema. To derive the ZF extensionality axiom, we first prove in TZF^- that $u = [z \,\Delta\, u]^u_z$ for all properties u. From this we may derive in T1 that, for all properties x and y, $\Box(\forall z)(z \,\Delta\, x \equiv z \,\Delta\, y) \supset x = y$. And from this plus the theorem that every instance of a pure L-determinate property is a pure L-determinate property, we derive

$$(\forall x, y)(x \text{ and } y \text{ are pure L-determinate properties} \supset ((\forall z)(z \text{ is a pure L-determinate property} \supset (z \,\Delta\, x \equiv z \,\Delta\, y)) \supset x = y)).$$

But given the contextual definition of the sentences of ZF in terms of Δ, =, and intensional abstraction, this sentence is just the expanded form of the ZF principle of extensionality:

$$(\forall x, y)((\forall z)(z \in x \equiv z \in y) \supset x = y).$$ *End of proof.*

Next take the axioms of von Neumann-Gödel-Bernays class theory, drop the axiom of extensionality, and rewrite the remaining axioms using Δ, =, and intensional abstraction.[24] TGB^- is the intensional logic obtained when these axioms are adjoined to T1.[25]

Metatheorem: Every sentence that is a theorem of von Neumann-Gödel-Bernays class theory is, given its contextual definition in terms of Δ, =, and intensional abstraction, a theorem of the intensional logic TGB$^-$ (i.e., for every set-theoretical sentence A, if $\vdash_{\mathrm{GB}} A$, then $\vdash_{\mathrm{TGB}^-} A$).[26]

The proof of this is analogous to the previous proof.

The intuitive content of the results is easy to state: Zermelo-Fraenkel set theory and von Neumann-Gödel-Bernays class theory have an alternate interpretation according to which they are just theories of pure L-determinate properties, a kind of property that forms a sub-universe within which there are no intensional distinctions.

31. Applied Set Theory Without Sets

I will now show how to give a no-class construction for a fairly elementary applied set theory which countenances empirical sets of particulars and of type 1 PRPs. The intuitive idea is that notation that previously had been interpreted as being about such empirical sets will now be introduced as an abbreviation for a longer property-theoretic locution that concerns the properties common to the elements of these empirical sets. (This way of treating set-theoretical notation is reminiscent of Russell's no-class construction of type-stratified set theory.) The construction, however, can be extended by analogy to more sophisticated applied set theories, including ones that countenance empirical sets of sets, etc. and that are fitted out with intensional and extensional abstraction operations.

$\mathscr{S}$ is a first-order language for elementary applied set theory. The primitive symbols of $\mathscr{S}$ are:

Logical operators:	$\&, \neg, \exists$
Predicates:	$=, \in, \Delta, F^1_1, \ldots, F^q_p$
Variables:	$x, y, z, \ldots$
	$\alpha_1, \alpha_2, \alpha_3, \ldots$
	$\beta_1, \beta_2, \beta_3, \ldots$
Punctuation:	(,).

Atomic formulas: $v_i = v_j$, $\alpha_i = \alpha_j$, $\beta_i = \beta_j$, $\alpha_i \in \alpha_j$, $v_i \in \beta_j$, $v_i \,\Delta\, v_j$, $F^j_i(v_1, \ldots, v_j)$. Let complex formulas be built up from these in the usual way. The variables $x, y, z, \ldots$ are to be thought of as ranging over particulars and type 1 PRPs; $\alpha_1, \alpha_2, \alpha_3, \ldots$, over pure sets; $\beta_1, \beta_2, \beta_3, \ldots$, over empirical sets of particulars and type 1 PRPs. And $F^1_1, \ldots, F^q_p$ are non-set-theoretic predicates. Now every sentence A of $\mathscr{S}$ can be contextually defined in L_ω with Δ as follows: substitute Δ for all occurrences of $\in$; replace all atomic formulas $\beta_i = \beta_j$ with $(\forall x)(x \,\Delta\, \beta_i \equiv x \,\Delta\, \beta_j)$; restrict all quantifers on pure set variables to pure L-determinate properties; replace all pure and empirical set variables with new distinct non-set variables. The result A^* is a sentence in L_ω with Δ. Then adopt the following contextual definition: A $\mathit{iff}_{\mathrm{df}}$ A^*. To see that this definition does the job, consider the theory $\mathrm{TZF}_{\mathrm{a}}^{-}$ which is just like TZF^{-} except that now $F_1^1, \ldots, F_p^q$ may occur in the axiom schemas. In $\mathrm{TZF}_{\mathrm{a}}^{-}$ we can derive, not only all the closures of the axioms of pure Zermelo-Fraenkel set theory, but also all closures of the following two axioms for the applied set theory:

(Extensionality)

$$(\forall x)(x \in \beta_i \equiv x \in \beta_j) \supset \beta_i = \beta_j$$

(Comprehension)

$$(\exists \beta_i)(\forall v)(v \in \beta_i \equiv (v \,\Delta\, u \,\&\, A))$$

where A is any formula of $\mathscr{S}$ in which β_i does not occur free. Hence, we have a no-class construction for not only pure first-order set theory but the applied first-order set theory as well.[27]

Summing up, we have seen that both pure and applied first-order set theories can be modeled within the first-order logic for the predication relation. Therefore, in view of the conclusion that the ontology of sets does not fall within our naturalistic ontology and the conclusion that set theory is not part of logic, there is simply no justification for positing the extraordinary abstract aggregates of set theory over and above PRPs and ordinary aggregates.

With this conclusion in hand I want to back up a bit. In the last three sections I have been operating under the assumption that set theory has at least a provisional role to play in mathematical matters. But now I want to challenge even that assumption, at least as it pertains to the analysis of numbers. For in the next chapter I will defend the thesis that in a proper construction of classical mathematics numbers should not even provisionally be identified with sets. Numbers should boldly be identified with properties.

6

Number

> It was Frege who first forced both philosophers and mathematicians to acknowledge the lack of any philosphical account of the nature and epistemological basis of mathematics. He himself constructed a complete system of philosophy of mathematics . . . . [T]he philosophical system, considered as a unitary theory, collapsed when . . . shown to be incapable of fulfillment . . . by Russell's discovery of the set-theoretic paradoxes. . . . [M]uch as we now owe to Frege . . ., it would now be impossible for anyone to consider himself a whole-hearted follower. . . .
>
> Michael Dummett
> *Elements of Intuitionism*

These excerpts express what appears to be the prevalent attitude toward logicism among leading contemporary philosophers of mathematics. Despite this, I am still inclined to hold a logicist position. In what follows I will employ the theory of PRPs to defend it. Along the way I will reply to the standard criticisms of logicism, none of which hits its mark in my opinion. I begin by considering logicism in the context of arithmetic. This after all was what Frege himself was concerned with, and it is here that the doctrine is most defensible.

32. A Neo-Fregean Analysis

Ask a practicing mathematician what the Peano postulates for number theory are. If he does not have a philosophical or historical ax to grind, in the majority of cases he will state the following:

(1) 0 is a natural number.
(2) Natural numbers have unique successors.
(3) 0 is not the successor of anything.
(4) If the successor of x = the successor of y, then $x = y$.

(5) For all properties z, if 0 has z and if each successor of anything having z itself has z, then every natural number has z.[1]

Indeed, this informal statement of the Peano postulates is given in book after book on number theory. Now it is clear that the most direct and natural formalization of (1)–(5) goes as follows:

(1′) $NN0$
(2′) $NNx \supset NNx'$
(3′) $\neg(\exists x)\, 0 = x'$
(4′) $x' = y' \supset x = y$
(5′) $(\forall z)((0 \mathbin{\Delta} z \mathbin{\&} (\forall x)(x \mathbin{\Delta} z \supset x' \mathbin{\Delta} z)) \supset (\forall x)(NNx \supset x \mathbin{\Delta} z))$.

These then are what I will call the Peano postulates. In doing so I believe I am being faithful to actual informal mathematical practice.

Consider the following neo-Fregean analysis of natural number:

$0 =_{\mathrm{df}}$ the property of being a property with no instances.

the successor of $x =_{\mathrm{df}}$ the property of being a property with one more instance than the instances of x.

x is a natural number *iff*$_{\mathrm{df}}$ x has each property z that is had both by 0 and by the successors of things that have z.

Symbolized in L_ω with Δ this neo-Fregean analysis becomes:

$$0 =_{\mathrm{df}} [\neg(\exists u)\, u \mathbin{\Delta} y]_y$$

$$x' =_{\mathrm{df}} [(\exists u)(u \mathbin{\Delta} x \mathbin{\&} (\exists v)(v \not\mathbin{\Delta} u \mathbin{\&} y \doteqdot [w \mathbin{\Delta} u \vee w = v]^{uv}_{w}))]^{x}_{y}$$

$$NNx \ \mathit{iff}_{\mathrm{df}}\ (\forall z)((0 \mathbin{\Delta} z \mathbin{\&} (\forall y)(y \mathbin{\Delta} z \supset y' \mathbin{\Delta} z)) \supset x \mathbin{\Delta} z)$$

where $y \doteqdot z$ *iff*$_{\mathrm{df}}$ $(\forall w)(w \mathbin{\Delta} y \equiv w \mathbin{\Delta} z)$. According to this analysis, natural numbers are fixed, purely logical objects, as logicists have thought. (They are not mere theoretical posits, as, e.g., Gödel thought.) This is one of the characteristic claims of logicism. Another characteristic claim of logicism is that Peano's postulates can be derived from principles of pure logic. And the following surprising little theorem results if the neo-Fregean definitions are adopted:

Theorem: Peano's postulates are theorems of the intensional logic T2 (i.e., if $\vdash_{\mathrm{PP}} A$, then $\vdash_{\mathrm{T2}} A$).[2]

Proof. Assume the definitions. Then (1′), (2′), and (5′) are immediate consequences of the axioms for quantifiers and truth-functional connectives. In addition, (3′) follows directly from the following instance of axiom $\mathscr{A}9$:

$$[\neg(\exists u)u \,\Delta\, y]_y \neq$$
$$[(\exists u)(u \,\Delta\, x \,\&\, (\exists v)(v \not\Delta\, u \,\&\, y \mathrel{\underset{\cdot}{=}} [w \,\Delta\, u \vee w = v]^{uv}_{w}))]^{x}_{y}.$$

Finally, the following is a theorem obtained by a few applications of axiom $\mathscr{A}10$ and an application of axiom $\mathscr{A}11$:

$$x = y \equiv [(\exists u)(u \,\Delta\, x \,\&\, (\exists v)(v \not\Delta\, u \,\&\, z \mathrel{\underset{\cdot}{=}} [w \,\Delta\, u \vee w = v]^{uv}_{w}))]^{x}_{z} =$$
$$[(\exists u)(u \,\Delta\, y \,\&\, (\exists v)(v \not\Delta\, u \,\&\, z \mathrel{\underset{\cdot}{=}} [w \,\Delta\, u \vee w = v]^{uv}_{w}))]^{y}_{z}.$$

(4′) follows immediately.

33. Reply to Criticisms

So far so good. Now let us see how the neo-Fregean definitions stand up against two of the standard criticisms of Frege's original definitions. (A third criticism, given by Charles Parsons, will be considered in §35.)

The easiest criticism to avoid is that of Robert Hambourger ('A Difficulty with the Frege-Russell Definition of Number'). The criticism is given in four steps. First, it is claimed that the following two propositions must both be true:

(a) There is at least one possible world in which the number 1 exists but in which some object that exists in the actual world does not exist.
(b) One and the same entity is the number 1 in each possible world; that is, it is not the case that one entity is the number 1 in one possible world while a different entity is the number 1 in another possible world. (p. 410)

Secondly, given the extensionality of sets, 'a set that exists in the actual world exists in a second possible world only if everything that belongs to it in the actual world exists in that second world' (p. 413). Thirdly, 'under the definition offered by Frege and Russell, ... 1 is the set of all unit sets' (p. 409).[3] Fourthly, it follows that, if (a) is true, then (b) is false.

This criticism is easy to avoid since it applies only to those analyses of the natural numbers that identify them with *extensional* entities, e.g., sets. However, the neo-Fregean analysis identifies natural numbers with *intensional* entities, namely, properties. According to the analysis, the natural numbers are fixed, purely logical objects. Although it is possible for their contingent instances to be different from what they actually are (for example, $[u \text{ is a president of USA in } 1980]_u \, \Delta$ the successor of $[\neg(\exists u)u \, \Delta \, y]_y$, and yet it is possible that $[u \text{ is a president of USA in } 1980]_u \, \not\Delta$ the successor of $[\neg(\exists u)u \, \Delta \, y]_y$), their fundamental logical relations necessarily remain unchanged and, hence, so do their mathematical relations. And that is all it takes to avoid Hambourger's criticism.

The next criticism of the logicist definitions is that of Paul Benacerraf ('What Numbers Could Not Be'). Before considering this criticism I must say more about how the neo-Fregean analysis dovetails with the theory of the logical structure of natural language. Consider the following English sentence:

There are 12 apostles.

The first thing to notice is that '12' occurs as a singular term, as the following intuitively valid argument shows:

There are 12 apostles.
$5 + 7 = 12$.
―――――――――――――
∴ There are $5 + 7$ apostles.

Next observe that in English 'There are n Fs' is well-formed if and only if 'The Fs are n' is also well-formed. (E.g., the following is perfectly good English. Question: 'How many is your party?' Answer: 'We are twelve.') Moreover, 'There are n Fs' and 'The Fs are n' seem to be synonymous. Finally, the following equinumerosity principle is intuitively valid:

There are exactly as many Fs as Gs *iff* for some number n, there are n Fs and there are n Gs.

and, more generally:

There are exactly as many things that have property x as there are things that have property y *iff* for some number n, there are n things that have property x and there are n things that have property y.

An adequate logical syntax for natural language should account for these elementary facts.

Now sentences such as 'There are twelve apostles' are not straightaway representable in first-order languages with intensional abstraction and predication. By contrast, sentences such as 'The apostles are twelve' are. In fact, the treatment of the copula arrived at in chapter 4 and the treatment of plurals[4] tentatively arrived at in chapter 5 lead directly to the following simple, indeed automatic, representation of 'The apostles are twelve':

$$\{v\colon Av\}\ \Delta\ 12$$

where $\{v\colon Av\}$ is an extensional abstract contextually defined in terms of Δ, and 12 is a singular term.[5] Therefore, the easiest way to bring, e.g., 'There are twelve apostles' within the scope of first-order logic is to treat it as a (meaning preserving) transformation from 'The apostles are twelve'. This way no new underlying logical structures need to be posited; first-order quantifier logic with predication and intensional abstraction suffices. But now consider the above equinumerosity principle. Given this principle and given the indicated treatment of 'There are twelve apostles' and 'The apostles are twelve', it follows that the singular term 'twelve' must be semantically correlated with a property (call it x) such that x's instances include *only* properties having twelve instances and, for *all* properties having twelve instances, x's instances include at least one property having those twelve instances. The simplest such property is a property whose instances are all and only properties having twelve instances. But this is exactly what the number twelve is defined to be in the neo-Fregean analysis. Therefore, given this analysis, we arrive at an analysis of natural number that easily accounts for all the data from natural logic cited above. In addition, the analysis achieves this without having to posit any new logical structures. And finally, it achieves this in such a way that the Peano postulates are derivable from the provably sound and complete logic T2, a logic that is justified quite independently of issues in philosophy of mathematics.[6]

With these conclusions in hand I am ready to consider Paul Benacerraf's influential criticism of the various analyses of natural number, including Fregean analyses. Omitting whatever implicit premises Benacerraf might have in mind, one may summarize the criticism by means of the following argument:

> There are many different things that, for all we know, the natural numbers could be.
>
> ---
>
> ∴ The natural numbers could not be any of them.[7]

Now although this argument is invalid,[8] its force is to point up a problem: since there are several non-equivalent candidate analyses of the natural numbers and elementary number-theoretic language, we need a rationale for selecting an analysis as the correct one. This is Benacerraf's challenge to us. I will try to meet it by sketching a rationale for selecting the neo-Fregean analysis.

There is no chasm separating elementary number-theoretic language from the idiom of cardinality that is built into the logical structure of natural language. Elementary number-theoretic language is part of natural language. Therefore, the best analysis of elementary number-theoretic language is the one that is part of (entailed by) the correct analysis of the logical syntax of the idiom of cardinality that is built into natural language.[9] For this reason, the problem raised by Benacerraf is a special case of the general problem of finding a rationale for selecting a theory of logical syntax for natural language as the correct one. Thus, the problem is a special case of Quine's indeterminacy problem in the theory of natural language.

What is this indeterminacy supposed to be? A careful analysis of Quine's skeptical attack shows, I believe, that it is at worst a fancy case of underdetermination (though Quine attempts to deny this in 'On the Reasons for the Indeterminacy of Translation'). And many commentators are beginning to see the matter this way. Now underdetermination is a problem that besets virtually all theories regardless of subject matter. Take virtually any subject matter and virtually any body of data concerning that subject matter. Typically there will be several candidate theories that provide acceptable accounts of the data. The rational way to decide among such competing theories is on grounds of naturalness, simplicity, and elegance. If these grounds are used elsewhere in theory to solve the problem of underdetermination, it would be an unreasonable use of a double standard to depart from this practice only in the case of the theory of logical syntax or the philosophy of mathematics.

Quite independently of issues in philosophy of mathematics I had already arrived at a theory of logical syntax that leads directly and almost automatically to the neo-Fregean analysis of natural number. This analysis serves to explain in a simple and natural way a variety of syntactic, semantic, and logical phenomena in natural language. In addition, it does so without having to posit any new logical structures. Finally, this is achieved in such a way that the Peano postulates are derivable from the previously arrived at logical theory. For these (and other) reasons, the neo-Fregean analysis seems simpler and more natural than its competitors. If indeed it is, then we are justified in identifying it as the correct analysis. And that, in my view, is the basis for a solution to the problem raised by Benacerraf's criticism.[10]

34. The Derivation of Mathematics from Logic

I have shown that the Peano postulates (1)–(5) stated above can be derived from principles of pure logic. How, then, from Peano's postulates does one go on to derive the rest of the elementary arithmetical truths expressed by sentences built up from $=$, NN, 0, and $'$? Consider an example. If $A(x)$ is the formula $NNx \; \& \; (x \neq 0 \supset (\exists y)(NNy \; \& \; x = y'))$, how from Peano's postulates does one infer that $(\forall x)((A(0) \; \& \; (\forall x)(A(x) \supset A(x'))) \supset (\forall x)(NNx \supset A(x)))$? In particular, how does one infer it from postulate (5), which says that, for all properties z, if 0 has z and if each successor of anything having z itself has z, then every natural number has z, i.e., $(\forall z)((0 \, \Delta \, z \; \& \; (\forall x)(x \, \Delta \, z \supset x' \, \Delta \, z)) \supset (\forall x)(NNx \supset x \, \Delta \, z))$? What the practicing mathematician typically would do (overtly or covertly) is to apply the following trivial validity:

> x has the property of being a natural number that is distinct from 0 only if x is the successor of some natural number *iff* x is a natural number that is distinct from 0 only if x is the successor of some natural number.

i.e.,

$$x \, \Delta \, [NNx \; \& \; (x \neq 0 \supset (\exists y)(NNy \; \& \; x = y'))]_x \\ \equiv (NNx \; \& \; (x \neq 0 \supset (\exists y)(NNy \; \& \; x = y'))).$$

Given this trivial validity, the arithmetical truth I wanted to derive follows immediately from postulate (5).

Now let me be very clear here. I am not talking about what the philosophically stern extensionalist logician would say ought to be done. I am talking about what a practicing mathematician in fact would typically do in making the inference from his own working property-theoretic statement of Peano's postulates, i.e., what his actual thought step would typically be as he reasons. I claim that what the practicing mathematician would typically do is to make the inference in question as a *purely logical* step. It would be perverse to insist that in making the inference he appeals to some further non-logical principle that he had omitted by oversight from his list of mathematical axioms (i.e., from his property-theoretic Peano postulates).

But notice that if we are given the single additional validity

$$x \,\Delta\, [NNx \,\&\, (x \neq 0 \supset (\exists y)(NNy \,\&\, x = y'))]_x \\ \equiv (NNx \,\&\, (x \neq 0 \supset (\exists y)(NNy \,\&\, x = y')))$$

then the complete theory for the structure $\langle NN, =, 0, ' \rangle$ is derivable from the pure logic T2.[11] One might argue against this that the above principle is not logically valid because, after all, closely related principles might somewhere down the line produce logical paradoxes. However, this is like arguing, e.g., that 'it is true that snow is white *iff* snow is white' is not logically valid because closely related sentences might give rise to the Epimenides paradox. Or even worse, that this is not oxygen because closely related gases might be noxious to humans. What is important here is that the particular principle in question is logically valid. And it unquestionably is.

Now I call formulas of the form

$$\langle v_1, \ldots, v_j \rangle \,\Delta\, [A]_{v_1 \ldots v_j} \equiv A(v_1, \ldots, v_j)$$

principles of predication. Consider the trivial principles of predication that are needed to derive from Peano's postulates classical number theory (with + and ·) and real and complex analysis.[12] As in the above case, it is clear that these few principles of predication are used in making purely logical inferences. They are not newly discovered non-logical principles that the mathematician forgot to include among his mathematical axioms. The particular principles of predication used are unquestionably logically valid. And so it is that from the pure logic T2 and a few additional validities classical mathematics can be derived.

35. Reply to Criticisms

Constructivism

The constructivist would object to the above derivation of classical mathematics on the grounds that some of the principles of predication used are non-constructive in character and, hence, that they are not valid. But a trap of sorts has been set for the constructivist. Recall the principle of predication discussed earlier:

$$x \,\Delta\, [NNx \,\&\, (x \neq 0 \supset (\exists y)(NNy \,\&\, x = y'))]_x$$
$$\equiv (NNx \,\&\, (x \neq 0 \supset (\exists y)(NNy \,\&\, x = y'))).$$

Since the range of all variables in the intensional abstract $[NNx \,\&\, (x \neq 0 \supset (\exists y)(NNy \,\&\, x = y'))]_x$ is restricted to natural numbers, this principle is the very paradigm of a constructive principle as conceived by the typical constructivist. Therefore, the constructivist has no choice but to accept it. Hence, if it is the case that when properly analysed NN, 0, and $'$ turn out to be non-constructive notions, then the constructivist is committed to accepting a non-constructive principle of predication. And if he is committed here, it is hard to see what grounds he could have for drawing the line when it comes to the classical theory of the real numbers. So in this way the consistency of the typical constructivist's philosophy of mathematics turns on an issue in philosophical analysis and the theory of logical syntax for natural language, namely, the issue of how various numerical constructions found in natural language are to be properly analysed. However, I have already given evidence for concluding that the neo-Fregean definitions are the right ones. And according to these definitions the natural numbers and the arithmetical operations are distinctly non-constructive. So unless the constructivist can meet the challenge either to discredit the evidence or to produce a better theory of the logical syntax of natural language (frequently the constructivist does not even seem to be aware of this challenge), it would appear that his philosophy of mathematics is not consistent.

Set Theory

Perhaps the most common objection to logicism these days is that the logicist construction of classical mathematics makes use of set theory and yet set theory is not part of logic. I agree that set theory is not part of logic; indeed, that was part of the thesis of the preceding chapter. But this fact is irrelevant to the logicist construction I am advocating, for this construction makes no use of set theory. Rather, the background theory is logic, specifically, first-order logic with intensional abstraction and predication.

Axiom of Infinity

The next criticism of logicism is that it needs to assume an axiom of infinity yet such an axiom lies outside the province of logic *per se*. While this criticism does tell against several logicist constructions (most notable of which is the type-theoretic construction in *Principia Mathematica*), it does not apply to the one proposed here. The various principles of predication that are employed in this construction are as close as I come to assuming an axiom of infinity. Consider, e.g., the principle of predication:

$$x \,\Delta\, [NNx]_x \equiv NNx.$$

To see that this trivial validity is not an axiom of infinity, adjoin it to elementary first-order extensional logic with identity. The existence of a property having infinitely many instances still cannot be inferred. It can be inferred only when appropriate auxiliary laws for NN—specifically, $NN0$ and $(\forall x)(x \neq \overbrace{x'''\cdots'}^{n})$, for all $n \geq 1$—are adjoined. The reason for this is that these auxiliary laws for NN, and not the principle of predication $x \,\Delta\, [NNx]_x \equiv NNx$, are what insure the existence of the infinitely many entities satisfying the predicate NN. Therefore, the auxiliary laws for NN are what yield the infinite ontological commitment. The principle of predication merely insures that the entities satisfying the predicate NN are indeed instances of the property that is expressed by NN. To be sure, given the neo-Fregean definitions, the auxiliary laws for NN can be derived in the intensional logic T2. But this is no embarrassment to logicism, for the infinite abstract ontology of intensional logic was already justified (chapter 1) quite independently of issues in philosophy of mathematics. Hence, unlike most logicist constructions, the neo-Fregean construction needs no special axiom of infinity, nor does it make any ontological commitments motivated by extra-logical considerations.

Incompleteness

The next criticism is based on Gödel's first incompleteness theorem.

The criticism is this: since, given Gödel's theorem, there exists no complete recursive axiomatization of number theory, not all truths of mathematics can be derived from logical validities; hence, logicism is false.

Notice, however, that this criticism goes through only when it is assumed that the validities have a complete recursive axiomatization. But this assumption seems to me to be false. For, given the logicist analysis of number, what Gödel's theorem shows is just that the validities have no complete recursive axiomatization.

This defense of logicism depends on how the concept of a logical validity is defined. I will take up this topic in §47, where the concept is defined formally. Given that definition, however, the defense of logicism is sustained. Although at this stage of the book it is not possible to go over this material in detail, I can say a few words about the underlying philosophical issue.

Informally speaking, the valid propositions are those that must be true in virtue of their logical form. What is the logical form of a proposition? Given the general framework of PRPs, there are two opposing views. On the first view the logical form of a proposition is merely the abstract shape of its decomposition tree as determined by the inverses of the type 2 fundamental logical operations (Conj, Neg, ...), where the particular identity of the various nodes in this decomposition tree is disregarded. The second view is just like the first except that the identity of the purely logical nodes (e.g., nodes occupied by identity, necessary equivalence, predication, etc.) is counted in. According to the first view of logical form, the valid propositions do have a complete recursive axiomatization. According to the second view, they do not. The second notion is, I believe, the right one. For observe that it is only on the second view that elementary truths involving identity and necessary equivalence—e.g., the propositions that $(\forall x)\, x = x$ and that $(\forall x)\, x \approx_{\mathrm{N}} x$—are counted as valid. However, if the purely logical relations identity and necessary equivalence are counted in, then surely there can be no rational grounds for not counting in the purely logical relation predication.[13] But when the predication relation is counted in, two important consequences follow immediately. First, given the neo-Fregean analysis of number, the truths of mathematics turn out to be validities. Second, there exists no complete recursive axiomatization of the validities.

Logicism in no way requires that there should be a complete recursive characterization of the validities; it requires only that the truths of mathematics should be validities.[14] And given the neo-Fregean definitions and the definition of validity, this is so.

Analyticity

While on the topic of validity, I should mention a criticism of logicism derived from Quine's attack on the notion of analyticity. Logicists claim that mathematics is analytic. The Quinean criticism is simply that the concept of analyticity is undefinable; hence, the logicist claim is not meaningful and, as such, is not true. However, the notion of analyticity can be rigorously defined within the theory of PRPs (see §47). Appropriately enough, analytic propositions are exactly those that are valid. Consequently, the truths of mathematics are analytic in a clearly defined sense. And this conclusion is arrived at without significantly distorting the original Kantian usage of the term 'analytic'.

The Failure to Find an Intuitive Complete System of Predication Principles

Consider the particular principles of predication that are used in the derivation of classical mathematics from Peano's postulates. Philosophers who would doubt these principles often do so as a result of the following faulty line of reasoning: the easiest syntactic generalization on these principles gives rise to an inconsistent system of logic; therefore, the principles themselves are called into doubt.

This line of reasoning is faulty, for it is based on the assumption that the easiest syntactic generalizations on sentences that express validities should lead to valid general principles of logic. This is an unjustified assumption. (Indeed, the tendency to make easy syntactic generalizations is partly what makes one so susceptible to the paradoxes in the first place.) To be sure, some parts of logic behave rather like this. When this is so, work goes very smoothly for logicians. However, given its essential incompleteness, the logic for the predication relation does not in general behave in this formally orderly way. In fact, there might be no intuitive complete system of predication principles, not even a very complicated one. Simply because a number of particular principles of predication are obviously valid, why should they not at the same time be full-blooded creatures of this incompleteness in logic in just the sense that they always defy syntactic generalization? Their resistance to syntactic generalization provides no more evidence for their invalidity than does the unprovability of an intuitively true "Gödel sentence" provide evidence for its falsehood.

To my knowledge no one has shown that an intuitive complete system of predication principles does not exist. In historical time the search for such a system is very young. There are several systems of logic that without unreasonable distortion may be viewed as generalizations on intuitively valid principles of predication. To find a more nearly perfect system requires finding appropriate features to generalize on. It would be no disaster, however, if there were none. General systems of logic are nice, but they are not required in order to have knowledge of particular validities. Logicism in no way depends on the existence of a companion general system of logic. The paradoxes indicate difficulties in the science of logic but not in logicism, for logicism is a doctrine in the philosophy of mathematics that concerns only the logical status of mathematics.

Large Numbers

A quite specific line of attack is that logicism is committed to the sort of unconditioned totality that leads to the paradoxes:

> ...[T]he application of numbers must be so wide that, if *all* concepts (or extensions of concepts) numerically equivalent to a concept F are members of N_xFx [i.e., the number of Fs], then it is by no means certain that N_xFx is not the sort of 'unconditioned totality' that leads to the paradoxes.
>
> ...[I]n the most natural systems of set theory, such as those based on Zermelo's axioms, the existence of ordinary equivalence classes is easily proved, while if anything at all falls under F, the non-existence of Frege's N_xFx follows. (p. 185, Charles Parsons, 'Frege's Theory of Number')

I will give four responses to this criticism. Since each is promising, I will not settle on any one of them.

(1) For a moment assume the whole of the Zermelo-style theory of properties, including the axioms that are used to prove Cantor's theorem and, in turn, the theorem that there exists no property that everything has. Even in this case, it appears that there is no difficulty in defining numbers in the neo-Fregean way. Reminiscent of Hambourger's problem, the problem here of a clash with Zermelo-based theories appears to arise only if natural numbers are identified with entities that must be picked out by reference to their

instances or members. The problem does not appear to arise if natural numbers are identified with entities that are picked out, not by reference to their instances or members, but rather by use of a canonical rigid designator such as an intensional abstract. Let me explain.

Consider an arbitrary natural number, e.g., 1. Suppose that 1 is indeed the property of being a property having one more instance than the instances of 0; i.e., suppose that $1 = [(\exists u)(u \, \Delta \, 0 \,\&\, (\exists v)(v \not\Delta \, u \,\&\, y \doteq [w \, \Delta \, u \vee w = v]^{uv}_{w}))]_{y}$. Assume further that every singleton property in fact has this property. Take the union of this property, i.e., the property of being an instance of an instance of 1. In a Zermelo-style theory of properties everything has this new property. However, in a Zermelo-style theory of properties it is also a theorem that there exists no property that everything has. Hence a contradiction. But notice that this argument makes use of an assumption, namely, the assumption that every singleton property has the property with which 1 has been identified. This assumption would hold if the extension of the predication relation were exactly what one would naively take it to be, i.e., if $\mathscr{G}(\bar{\Delta}) = \{xy \in \mathscr{D} : x \in \mathscr{G}(y)\}$. But the Zermelo-style resolution of the paradoxes in property theory modifies things just here. I see no way to derive the above assumption in a Zermelo-style theory of properties. Thus, I conjecture that the neo-Fregean analysis need not clash with Zermelo-style theories.[15]

(2) Even if I should be wrong about the foregoing, there is a way to preserve the neo-Fregean analysis within a more or less Zermelo-based theory of properties, namely, within the von Neumann-style theory.[16] The only thing that needs to be changed is the definition of *NN*. The new definition of *NN* is then derived from Dedekind's definition, not Frege's. Finite and transfinite arithmetic are still derivable. In addition, every singleton property that has properties at all has the property with which 1 is identified. (Likewise for doubleton properties and the number 2; and so on.) This is possible because natural numbers now are proper objects (i.e., unsafe properties). Hence, natural numbers are equinumerous with, e.g., $[x = x]_x$, a property that *all* safe objects have. So the threatened clash between the neo-Fregean analysis of number and this more or less Zermelo-based theory is avoided.

(3) A third way to respond to Parsons' criticism within the setting of a Zermelo-style theory goes as follows. Recall the reso-

lution of the paradoxes in property theory modifies things just here. sidered in §26. (Parsons himself favors such a resolution in 'The Liar Paradox'.) According to this resolution, in all contexts of speech and thought there is an implicit limitation on the universe of discourse. If this line is adopted, then it should be applied to numerical expressions and concepts as well. Hence, in the neo-Fregean definitions all variables within intensional abstracts ought to be grounded, i.e., they ought to be restricted in their range to a given antecedently fixed universe of discourse u. So, for example, relative to a context in which u is the implicit universe of discourse, the neo-Fregean definition of zero might go as follows:

$$0_u =_{\mathrm{df}} [x \Delta u \,\&\, \neg(\exists y)(y \Delta u \,\&\, y \Delta x)]^u_x.$$

Since in the context of arithmetic (finite or transfinite) the identity of u plays no special role (as long as u is sufficiently large for the purposes at hand), explicit occurrences of u may be suppressed.

(4) A fourth response to Parsons' attack is just to reject the Zermelo-based logics for the predication relation. After all, the Zermelo-based logics are not all that natural. And there are rather natural alternatives. For example, the logic for the predication relation that is based on Quine's NF has a variety of attractive features. E.g., according to it everything has the property of being something; i.e., the concept of a thing applies to all things. Even if one has reservations about the existence of a universal set, these reservations do not obviously carry over to properties or to concepts. There seems nothing absurd in holding that everything has the property of being something, that the concept of a thing applies to all things, including itself. This picture of properties and concepts is in fact very much like the view that Gödel himself arrived at.[17] Now in the logic for the predication relation that is based on NF, not only does everything have the property of being a thing, but also every property that is a property of exactly one thing has the property of being a property having one more instance than the properties having no instances; e.g., $(\forall x)[y = x]^x_y \Delta 1$. In view of this, Zermelo-based theories can hardly be used as the acid test for the correctness of the neo-Fregean analysis of number.

Given the foregoing options, it would seem, then, that Parsons' criticism of Frege's original analysis does not undercut the neo-Fregean analysis.

36. Epistemological Issues

We come finally to the epistemological worries about logicism. Consider the principle of predication that was used to derive from logic the complete theory for $=$, NN, 0, and $'$:

> x has the property of being a natural number that is distinct from 0 only if x is the successor of some natural number *iff* x is a natural number that is distinct from 0 only if x is the successor of some natural number.

i.e.,

$$x \,\Delta\, [NNx \,\&\, (x \neq 0 \supset (\exists y)(NNy \,\&\, x = y'))]_x$$
$$\equiv (NNx \,\&\, (x \neq 0 \supset (\exists y)(NNy \,\&\, x = y'))).$$

How do we know this? The answer is that we know it in the same way that we know, e.g., that x has the property of being Socrates *iff* x is Socrates. We know it in the same way that we know elementary axioms of first-order quantifier logic. We know it in the same way that we know that the various instances of modus ponens are valid. We know it in the same way we know, e.g., the theorems of T1 or T2, say, the theorem that $x \,\Delta\, [NNx]_x \vee x \not\Delta\, [NNx]_x$. These truths of logic are just as obvious, trivial, absurd to doubt, etc., as the principle of predication mentioned above. And as far as I can tell the same thing goes for every single one of the principles of predication that is used in the derivation of classical mathematics from Peano's postulates (which are themselves derivable from the logic T2).

The Need for a Complete Epistemological Account

The logicist has often been criticized for not having kept his promise to provide an account of mathematical knowledge. But this is wrong. The logicist does provide an account of mathematical knowledge; it is just not complete. The account, as far as it goes, is this. Elementary or complex mathematical truths are identical to complex logical validities. Thus, knowledge of mathematical truths is knowledge of complex validities. Hence, mathematical knowledge has the very same explanation—whatever it is— as does knowledge of complex validities. As far as I am concerned this is all the logicist needs to say, for the logicist is not required to give an account of how we come to have knowledge of complex validities. That is a general topic in epistemology, not philosophy of mathematics. It is a topic about which the logicist *qua* logicist should say nothing in particular except that knowledge of complex validities is not dependent on some further special kind of mathematical knowledge such as non-logical, pure mathematical intuition. The demand that the logicist provide a complete epistemological account is based on a confusion about the relationship between the special hybrid areas in philosophy such as philosophy of mathematics, philosophy of science, philosophy of law, etc., and the primary general areas of philosophy, namely, epistemology, metaphysics, and value theory. A successful theory in one of the special hybrid areas is one that can be integrated with successful general theories in the primary areas.

An Integrated Epistemological Account

My last remark brings me to a further epistemological worry about logicism, namely, the worry that the logicist account of mathematical truth is such that it cannot be integrated into any successful epistemology. This worry is voiced by Paul Benacerraf in his paper 'Mathematical Truth':

> ...two quite distinct kinds of concerns have separately motivated accounts of the nature of mathematical truth: (1) the concern for having a homogeneous semantical theory in which semantics for the propositions of mathematics parallel the semantics for the rest of language, and (2) the concern that an account of mathematical truth mesh with a reasonable epistemology. It will be my general thesis that all accounts of the concept of mathematical truth can be identified with serving one or another of these masters *at the expense of the other*. (p. 661)

Now the logicist theory of mathematical truth that I have given above yields the kind of homogeneous semantics mentioned in (1). So does this logicist theory mesh with a reasonable epistemology? I will try to show that it does. But I will not try to show this categorically; rather, I will attempt to establish my point relative to the plausible standard that Benacerraf suggests in his paper. Specifically, Hilbert's theory is identified there as the paradigm of a theory of mathematical truth that meshes with a reasonable epistemology.[18] Surely no one will object to this standard.

My argument is really very simple. On Hilbert's view it is a precondition of our having knowledge of complex logical and mathematical truths that we should have knowledge of the axioms and rules of elementary extensional logic with identity. Hilbert, however, provides no theory of how we come to have the latter kind of knowledge. Still, his theory presumably meshes with a reasonable epistemology. Now consider the axioms and rules of the logic for L_ω and consider the modest principles of predication that suffice for the derivation of classical mathematics from these axioms and rules. On the face of it, we seem to know these elementary axioms, principles, and rules in exactly the same way that we know the axioms and rules of elementary extensional logic with identity. As far as a reasonable epistemological account is concerned, then, knowledge of these elementary truths of intensional logic with predication would on the face of it appear to be quite on a par with knowledge of elementary truths of extensional logic with identity. A reasonable account of one would on the face of it appear to be easily adapted so as to provide an account of the other. Therefore, those who are skeptical are obliged to produce a reason for doubting that this is indeed the case. Without such reason we may safely conclude that, if there is a reasonable epistemological account of our knowledge of elementary extensional logic with identity, then there is a reasonable epistemological account of our knowledge of the above elementary principles of intensional logic with predication. Given this conclusion the rest of the argument is easy. For on Hilbert's view (like Frege's view before it), if there is a reasonable epistemological account of our knowledge of the premises and rules that we use in giving a proof, then there is a reasonable epistemological account of our knowledge of what we have proved. However, given the neo-Fregean analysis of number, we can prove the theorems of classical mathematics using known premises and rules. And given the previous conclusion, our knowledge of these premises and rules is assured of having a reasonable epistemological account. It follows, therefore, that relative to the plausible standard of reasonableness indicated earlier, our knowledge of classical mathematics is assured of having a reasonable epistemological account that meshes with the logicist theory of mathematical truth.

Now someone might object that, whereas the logicist is faced with the question of how he knows that his premises and rules are consistent, the follower of Hilbert is faced with no analogous question, for elementary extensional logic with identity has been proved consistent. In addition, someone might object that, whereas the logicist is faced with the question of how to account for the acquisition of new knowledge of the infinite not derivable from his present premises and rules (e.g., knowledge of the axiom of choice, knowledge of the continuum hypothesis or its negation, etc.), the follower of Hilbert is not faced with analogous problems. However, both of these objections, I would suggest, result from a failure to appreciate the force of Gödel's two incompleteness theorems.

Consider the first objection. The proof of the consistency of elementary extensional logic with identity is given within a background theory that is stronger than the original theory. In fact, the background theory includes elementary extensional logic with identity and concatenation. Moreover, this concatenation theory (as normally formulated) is equivalent to first-order arithmetic. But Gödel's second incompleteness theorem states that the consistency of first-order arithmetic cannot be proved without appealing to a still stronger background theory. And so it goes. Thus, if one is really in doubt about the consistency of elementary extensional logic with identity, it is difficult to understand why these consistency proofs should resolve the doubt. (Kleene reports that Tarski, when asked whether he felt more secure about classical mathematics from Gentzen's proof, replied, 'Yes, by an epsilon.') With regard to the question of how consistency can ever be really known, the follower of Hilbert is in no better position than the logicist. To be sure, the stronger a theory the greater its risk of inconsistency. But differences here are only in degree, not in kind. So regarding actual epistemological security, the logic espoused by the follower of Hilbert differs from the one espoused by the logicist only in degree.

Next consider the second objection. Gödel's first incompleteness theorem is that there is no complete axiomatization of first-order arithmetic. Therefore, regardless of how rich one's axiomatization of arithmetic is, it will always be possible to discover new elementary truths about the natural numbers that cannot be proved from those axioms. How is such new knowledge of the finite acquired? The follower of Hilbert would seem to be obliged to answer this question. However, with regard to the possibility of finding an answer that is compatible with a reasonable epistemology, this question seems to be analogous to the question facing the logicist. Thus, against these two objections, it still appears that, concerning its compatibility with a reasonable epistemology, logicism is not essentially worse off than Hilbert's theory.

But how do we know elementary logical validities? This question, like the question of how we know complex validities, is a general question in epistemology. It has no special dependence upon the problems in philosophy of mathematics that logicism is designed to solve. On this question the logicist and his competitors, including the followers of Hilbert, are all in the same boat. My own inclination is to think that some kind of rationalistic answer to this question is the only reasonable one. My reason, put bluntly, is that if someone is not endowed with powers of reason sufficient for *a priori* knowledge of elementary validities, he will alas not be intelligent enough to learn them *a posteriori* either.

Reduction and Degrees of Certainty

There are other epistemological criticisms of logicism, e.g., those of Poincaré, Wittgenstein, and Charles Parsons. In *Mathematical Knowledge* Mark Steiner cogently rebuts these criticisms. He goes on to defend the popular anti-logicist position that natural numbers are irreducible objects. The defense is itself epistemological, and it proceeds as follows. Arithmetic unreduced is more certain than arithmetic reduced *à la* logicism; at the same time, logicist reductions of arithmetic do not make any improvements on arithmetic unreduced; therefore, the best unified mathematical theory is one that keeps arithmetic in its unreduced form, rejecting all logicist reductions.

I find this argument dubious. But before I state my objection, let us take note of the view of reduction that is at work here. For if this view were correct, then my objection would be unfounded:

> One might wonder whether reduction is then ever possible, since all reductions seem to reduce a weak but more certain theory (Boyle's and Charles' laws) to a stronger but less certain theory (molecular theory). Such an objection would overlook a cardinal difference: bona-fide reductions effect changes that improve the original theories. They explain why the originals fail to be universally true. (p. 86, Steiner, *Mathematical Knowledge*)

I find this standard for what it takes to be a justified reduction far too high, for its leaves out the mundane cases of justified reduction. Consider the following mini-reduction:

> This pencil weighs more than 1 gram. ⇒
> The spatially present rigid collection of molecules that is producing these tactile sensations weighs more than 1 gram.

This reduction does not meet the above standard for two reasons. First, it is not clear that the mundane theory in any relevant sense fails to be universally true, and moreover, even if it should fail to be universally true, the explanation would be quite independent of the reduction. Secondly, the reduction does not *improve* upon the original mundane theory. There is nothing in the original theory that needs improving; it is just fine as it stands. And yet the reduction is as justified (this is not to say that it is as interesting) as any known reduction. The identification of this pencil with the spatially present rigid collection of molecules that is producing these tactile sensations is easily justified as follows. The molecular theory of perception provides us with good reasons for holding that space is populated widely by molecules and, in particular, that a spatially present rigid collection of molecules is producing these tactile sensations. Now consider the "dualist" theory that everyday material bodies inhabit the very same places as rigid collections of molecules. This theory is not defective in its observational content—that is not its flaw. What is wrong with it is that it is very uneconomical and, at the same time, it leaves unexplained the fact that so many properties of any given material body are the same as the properties of the rigid collection of molecules that inhabits the same place (e.g., it leaves unexplained why so many properties of this pencil are the same as the properties of the spatially present rigid collection of molecules that is producing these sensations). The reduction of everyday material bodies to rigid collections of molecules removes both of these defects. This alone is enough to justify the reduction. And this reduction is justified despite the fact that the reduced theory that this molecular collection weighs more than 1 gram makes no improvement on the unreduced theory and despite the fact that the unreduced theory (i.e., the everyday-body theory) is more certain than the reduced theory (i.e., the molecular-collection theory).

But look at the parallelism between the reduction of everyday bodies to rigid collections of molecules and the logicist reduction of natural numbers to properties. We have good logico-linguistic reasons (chapter 1) for holding that property theory is built into the logical syntax of natural language. And we have good logico-linguistic reasons (§33) for holding that a certain purely logical property is what corresponds semantically to a numerical adjective (e.g., '12') in such natural language sentences as 'There are 12 apostles'. Consider the "dualist" theory that what corresponds semantically to such numerical adjectives are different from the entities (i.e., the natural numbers) that correspond semantically to the typographically identical numerical expressions in the language of the science of arithmetic. What is wrong with this dualist theory is that it is uneconomical and, at the same time, it leaves unexplained the fact that so many properties of the two kinds of entities are the same. Reducing natural numbers to the indicated purely logical properties removes both of these defects. Now this justification parallels perfectly the one given for the everyday-body/molecular-collection reduction. Therefore, since the one is justified, so is the other. And this is so despite the fact that the reduction makes no improvement on any particular laws of arithmetic. Further, this would be so even if the unreduced theory (i.e., the laws of arithmetic) were more certain than the reduced theory (i.e., the laws of intensional logic with predication that are used in the reduction). And this concludes my argument.

Is the unreduced theory really more certain than the reduced theory? Many philosophers of mathematics today assume that it is, and they then use this assumption to upbraid the logicist for not providing mathematics with a new, epistemologically more secure footing than it had previously. I want to make a few comments against this popular position. To my mind the basic laws of arithmetic are not more certain than the associated laws of intensional logic with predication. After all, the latter laws are highly compelling. However, there is a deeper point here that many of the epistemological critics of logicism appear to have forgotten. Granted, many of us feel quite certain about the basic laws of arithmetic. Yet in reply to, say, someone who is seriously in doubt about the objective existence of natural numbers (e.g., a classical nominalist, a conventionalist, or an ontological relativist such as Carnap), what do we really have to say by way of justification for these laws?

First, there is the naive realist's commonsense reply that arithmetic is just one of the special sciences (i.e., one of the several epistemologically justified special disciplines), that its subject matter is the natural numbers, and that these laws are its first principles. But this reply hardly suffices, for it merely asserts the very sort of thing that is being challenged. Secondly, there is the intuitionistic reply that we know the basic laws of arithmetic by some special faculty of mathematical intuition.[19] But this will not do, for the existence of such a special faculty is as much in doubt as arithmetic itself. Thirdly, there is

Quine's reply that arithmetic is needed in the empirical sciences and, therefore, that it is justified by the empirical sciences taken as a whole. Quine's reply might be enough to overcome doubts about the objective existence of natural numbers. However, I think that we can do much better. For given Quine's view, arithmetic cannot be more certain than the empirical sciences. But this just seems false. Fourthly, there is Gödel's reply, which is an amalgamation of the intuitionistic reply and Quine's reply. Although Gödel's reply does not posit a faculty of mathematical intuition as an ultimate authority, it takes mathematical intuitions, collectively, to be a special body of empirical data to be explained (in part) by the associated special science of mathematics. According to Gödel, the best explanation of a (sound) mathematical intuition is that it is a special kind of perception of mathematical reality; our mathematical theories are just the best known systematization of these perceptions in much the same way that our physical theories are the best known systematization of our sense perceptions. Thus, mathematical theories are justified in much the same way that physical theories are justified. Although Gödel's reply, like Quine's reply, has certain virtues, it is not decisive. For it is not clear that Gödel's explanation of mathematical intuition is the best one.[20]

Finally, we come to the logicist's reply to doubts about the objective existence of natural numbers. In contrast to Gödel, the logicist need not appeal to mathematical intuition. Instead, his reply is that, upon analysis, the natural numbers are seen to be purely logical objects whose existence is independently justified by logic. Specifically, natural numbers are properties, and the ontology of properties, relations, and propositions is required in the best formulation of logic. In the same vein, the basic laws of arithmetic are, upon analysis, seen to be laws of logic. As such, they are justified in the way that laws of logic are usually justified. (As before, the logicist *qua* logicist is under no special obligation to say how our knowledge of laws of logic is justified.) Previously, skeptics in the philosophy of mathematics felt free to doubt such things as the existence of numbers. But now we see that such doubts are tantamount to doubting laws of logic. For this reason, if those who are skeptical about the foundations of mathematics are not careful, their doubts might compromise their commitment to logic.

By virtue of his ability to reply to doubts regarding the *foundations* of arithmetic (doubts about the existence of natural numbers are merely representative of a wide range of doubts that can be met in this way), the logicist provides a kind of epistemological justification not available under the competing philosophies of mathematics. And this justification does not depend on any epistemological resources not already employed in these competing philosophies. All that is required in addition is the neo-Fregean analysis of natural number.

How does one know that the neo-Fregean analysis of number is right? In the same way that one usually comes to know complex, informative definitions, namely, by having a theoretical justification. In the case of the neo-Fregean definitions, the justification is highly theoretical. In fact, it is pretty much the one laid out in the course of this chapter.

There is one more way to attempt to refute logicism epistemologically; it goes as follows. There are numerous arithmetic truths that seem obvious to the naive eye. Yet their neo-Fregean counterparts not only fail to be obvious to the naive eye but in fact are so complex that they almost defy understanding. This shows that the unreduced arithmetic truths are more certain than their neo-Fregean counterparts. And this, in turn, shows that the neo-Fregean analysis must be mistaken. For if it were correct, the unreduced arithmetic truths and their neo-Fregean counterparts would have to have the same degree of certainty.

To argue in this way, however, is to show that one has forgotten the paradox of analysis. In fact, the intensional logic developed for resolving the paradox of analysis (§20) nips this last little argument in the bud.

In this part of the work we have seen that the predication relation is what thrusts logic into incompleteness and indeed threatens it with paradox. And yet the predication relation also gives logic a great deal of power: the predication relation can claim responsibility for bringing all of classical mathematics into the domain of logic. This, however, is not the only way in which our conception of logic must be expanded. By the end of part III, it will have been expanded in several other surprising ways.

Part III

Unification

7

Extensionality and Meaning

Up to this point in my investigation of properties, relations, and propositions I have taken a free and easy approach toward three issues: the relation between extensional and intensional logic, the relation between the semantic theory of truth and the theory of meaning, and the relation between the two traditional conceptions of PRPs, conceptions 1 and 2. First, I have formulated intensional logic as if it were an autonomous subject above and beyond extensional logic. Secondly, I have characterized the semantics for intensional language by means of a Tarski-style theory of truth, and in so doing I have proceeded as if the theory of truth were an autonomous subject independent of the theory of meaning. Finally, I have developed the two traditional conceptions of PRPs side by side as if each one were fully correct and autonomous from the other. Taking these steps considerably simplified the investigation of several highly complex problems. But the three issues must be confronted directly before a fully satisfactory treatment of PRPs can be attained.

My goal will be unification: a construction of intensional logic within an extensional logic; a derivation of the theory of truth from a unified semantic theory based on a single meaning relation, and finally, a synthesis of the two traditional conceptions of PRPs within a theory of qualities and concepts.

In the classical tradition of Frege and Russell, the first two issues—the relation between extensional and intensional logic and the relation between the theory of truth and the theory of meaning—are treated in tandem. I will follow the same practice in this chapter. In the following chapter, which is more metaphysical than linguistic in character, I will discuss the relation between the two traditional conceptions of PRPs. The resulting synthesis will then be applied to a number of outstanding problems in modern philosophy. These applications comprise the final two chapters.

37. The Thesis of Extensionality

Why is there intensionality in language? Why, in addition to extensional language, for which Leibniz's law and the substitutivity of equivalent formulas are valid, is there any language for which these laws are invalid?

In chapter 1 we took a preliminary look at these questions. When one looks more closely, one may take either of two fundamentally different attitudes toward this question. According to the first, it is simply assumed that there really are violations of the laws of extensional logic. We are to take it as a fact of life that the extensional and the intensional are two primitively different kinds of linguistic forms having different kinds of logic. According to this attitude, one should not attempt to explain away intensionality in language. Rather, one should simply codify both kinds of logic, extensional and intensional, more or less as they appear to the naive eye. This attitude toward intensionality in language underlies the now dominant movement in modal logic originated by C. I. Lewis and carried on by Carnap in *Meaning and Necessity*[1] and, more recently, by Hintikka, Kripke, and their school. This attitude is also held by Montague in his 'Universal Grammar' and 'The Proper Treatment of Quantification in English'. And, of course, this is the attitude originally held by Russell in the first edition of *Principia Mathematica*, where we find the first comprehensive formal intensional logic.[2]

According to the second attitude toward intensionality in language, there are no genuine violations of the laws of extensional logic; all *prima facie* intensional phenomena are surface phenomena that can be explained away. All language, and all logic, is at bottom extensional. This attitude is, of course, the one held by Frege in 'Über Sinn und Bedeutung'. There Frege outlined a theory, subsequently formalized by Church, by means of which all *prima facie* intensional phenomena can be explained away. According to this theory, all *prima facie* deviations from extensional logic are produced by the fact that in the problematic contexts certain expressions do not name what they usually name; instead they name the intensional entities that they usually express. Thus, for Frege and Church, there is no genuinely intensional language; when *prima facie* intensional language is properly analysed, it turns out to be extensional language concerning intensional entities.

The attitude that intensionality in language is only an apparent phenomenon is also evident in Carnap's thesis of extensionality, which he advanced and defended in *The Logical Structure of the World* and *The Logical Syntax of Language* (only to abandon it in *Meaning and Necessity*). The thesis was the subject of lively debate from Wittgenstein's *Tractatus* to Carnap's *The Logical Syntax of Language*. Since then, however, it has been largely neglected, no doubt as a result of the prevalence of the first attitude, i.e., the one Carnap took in *Meaning and Necessity*. In *The Logical Syntax of Language* Carnap formulates the thesis of extensionality as follows:

> *a universal language of science may be extensional*; or, more exactly: for every given intensional language S_1, an extensional language S_2 may be constructed such that S_1 can be translated into S_2. (p. 245)

The truth of the thesis of extensionality would immediately suggest an account of intensionality in language. An especially perspicuous statement of this account may be given when one appeals to the theory of transformational grammar. If, as the thesis of extensionality insures, every intensional sentence can be translated into an extensional sentence, then every intensional sentence can be treated as the result of applying a meaning-preserving transformation to the original extensional sentence. The deep structure of language would be fully extensional, and intensionality in language would be a mere surface phenomenon. Such an account of intensionality in language would be quite general; indeed, Frege's account could be viewed as a special case of it.

In logical theory there are two competing general methodologies, one liberal and one conservative, and the conflict between the two attitudes toward intensionality in language may be viewed as an instance of this methodological conflict. According to the liberal methodology, when one attempts to extend the scope of logical theory to uncharted territory, one should feel free to adopt a pragmatic approach. Specifically, one should not feel constrained to formulate the laws of logic in such a way that they exhibit maximum generality. All that is required is that the new theory should work as a self-contained whole. According to the conservative methodology, by contrast, one is constrained to formulate the laws of logic in such a way that they do exhibit maximum generality: if a law is valid for a natural fragment of the language under consideration, then it should, if at all possible, be valid for the entire language. It is not enough that a logical theory should work as a self-contained whole.[3] There is in addition a demand for generality.

This methodological conflict shows up in numerous places in logical theory. For example, it is seen in the two classical theories of definite descriptions, Frege's and Russell's. On Frege's theory, definite descriptions are the result of applying an unanalysed name-forming operator (e.g., 'the α such that') to open-sentences (e.g., '$A\alpha$').[4] Sentences containing definite descriptions are then treated as if the definite descriptions were full-fledged names. However, in view of the phenomenon of vacuous definite descriptions, Frege's theory has the effect of overturning certain logical laws that held prior to the introduction of definite descriptions, namely, existential generalization and the law of the excluded middle. To the extent that Frege is willing thus to put limits on the syntactic domain in which these laws are valid, he may be viewed as a proponent of the liberal methodology, which does not demand maximum generality. In contrast to Frege's theory of definite descriptions, Russell's theory does not have the effect of overturning any logical laws that held prior to the introduction of definite descriptions. For on Russell's theory, sentences containing definite descriptions arise from the application of a meaning-preserving transformation to sentences that, ultimately, contain no definite descriptions. Hence, all logical laws that held prior to the introduction of definite descriptions—including the law of the excluded middle and existential generalization—hold for the enlarged language which contains definite descriptions. Thus, with regard to definite descriptions Russell's theory tends to maximize generality in the laws of logic. Inasmuch as this was his goal, Russell may be viewed as a proponent of the conservative methodology, which demands maximum generality.

Our present concern is the treatment of intensionality in language. The Lewis-style modal logician begins his study with a standard extensional language. To this language he adjoins new primitive operators $\Box$ and $\Diamond$, thereby obtaining an enlarged class of well-formed formulas. Since $\Box$ and $\Diamond$ are taken as primitive, the logical syntax of each new formula is identified with its surface syntax. As a result, the new formulas must be treated as genuine violators of the laws of extensional logic, i.e., laws that hold for the language to which the modal operators are adjoined. However, this loss of generality makes no difference to the Lewis-style modal logician. All that matters to him is that the logic for the enlarged language can be characterized in terms of his logical syntax. In this, the Lewis-style modal logician may be viewed as a proponent of the liberal methodology. By contrast, on the treatment of intensionality associated with the thesis of extensionality, all *prima facie* intensional expressions arise from the application of meaning-preserving transformations to expressions that ultimately conform to the laws of extensional logic. Hence, the laws that hold in the original language also hold in the enlarged language, generally. Thus, this treatment of intensionality promotes the maximum generally demanded by the conservative methodology.

This methodological conflict over the treatment of intensionality also shows up in the theories held by Russell and Frege. This time, however, Russell, not Frege, is the one representing the liberal methodology, for he posits primitive intensional forms as well as primitive extensional forms. And Frege, not Russell, is the one representing the conservative methodology, for he advocates a fully extensional logic for *prima facie* intensional language.[5] Other things being equal, methodological vacillations like these ought to be avoided.

It cannot be denied that the liberal methodology has led to a number of valuable advances in logic. Indeed, whereas the conservative approach often bogs down in the face of demanding constraints, the liberal methodology is free and easy. And inasmuch as it frees logical structure to reflect surface grammatical forms, it removes one of the barriers to a familiar and natural logical syntax. These reasons justify the liberal methodology at least as a short-term strategy. However, as a long-term strategy, only the conservative methodology guarantees the kind of maximally general theory that is the highest ideal in science. So it too should be explored.

Similarities among intensional expressions in natural language lead us to the thesis, defended in §§6–9, that there is a single intensional abstraction operation—represented by the bracket notation $[\ ]_\alpha$—that underlies all apparent intensionality in language. Now suppose that following the liberal methodology one treats this intensional abstraction operation as if it were primitive. (This is the strategy followed in chapter 2.) One would then arrive at an intensional logic that is highly successful at least on its own terms. To stop here would be to adopt the liberal methodology. Can this approach to intensional logic be brought into line with the conservative methodology, which calls for a logic that at bottom is fully extensional? That is, in accordance with the thesis of extensionality, is there a way to treat sentences containing intensional abstracts as the outcome of applying a meaning-preserving transformation to sentences that ultimately are extensional? Can intensional abstracts be defined, directly or contextually, within extensional language?

According to §8, intensional abstracts are semantically correlated, not with linguistic entities, but instead with intensional entities. In view of this, it would be unwise to base an extensional definition of intensional abstraction upon any of the nominalistic approaches (such as the one that Carnap offered in his original defense of the thesis of extensionality). It would also be unwise to pursue Frege's theory, for his approach to intensionality runs into troubles on such matters as Davidson's finite learnability requirement (desideratum 13; see §8) and quantifying-in (desideratum 5; see §11).[6]

Despite troubles in Frege's approach, one can draw inspiration from his informal doctrine that all *prima facie* intensional language is no more than extensional language about intensional entities. According to §8, an intensional abstract is semantically correlated with the intensional entity that the abstracted formula expresses. The key to giving an extensional definition of intensional abstraction, then, is to find a way to give extensional descriptions of the intensional entities semantically correlated with intensional abstracts. The algebraic theory of intensional entities is the crucial ingredient. On the resulting analysis, an intensional abstract would be treated as a transformation from a structural description of an intensional entity; further, this structural description would be stated in terms of the fundamental algebraic logical operations (conjunction, negation, existential generalization, etc.), and finally, the syntactic structure of the intensional abstract would stand in an easy one-one correlation to the syntactic structure of the structural description.

Some illustrations will make plain how this works. For example, 'the proposition that there exists an x such that Fx' is transformed from the (structural) definite description 'the proposition that is the existential generalization of the property F-ness'; 'the property of being an x such that Fx and Gx' is transformed from the definite description 'the property that is the conjunction of the properties F-ness and G-ness'; and so on for more and more complex intensional abstracts. (Both here and below I use the terms 'F-ness' and 'G-ness' for heuristic purposes only. Primitive property and relation names (e.g., 'red', 'love', etc.) would take over their role in the final analysis.)

Notice in these examples that when the predicate 'F' occurs within an intensional abstract, it does not actually occur as a predicate. Instead, it occurs in effect as a name, for it is correlated via the transformation to an occurrence of the name 'F-ness'. And this occurrence of the name 'F-ness' names the very intensional entity that is the meaning of the predicate 'F'. Thus, the transformation has the effect of giving us extensional occurrences of names that name the meanings of the predicates to which they are correlated. Much the same thing goes for *prima facie* occurrences of formulas within intensional abstracts. When a formula seems to occur within an intensional abstract, what one really has is an expression that is correlated via the transformation to an extensional occurrence of a definite description of the intensional entity that is the meaning of the formula. Hence, in general, the transformation has the effect of giving us extensional occurrences of singular terms that denote the meanings of the predicates and formulas to which they are correlated. This, of course, is reminiscent of certain aspects of the higher-order theories both of Frege and of the Russell of the first edition of *Principia Mathematica*.

With these preliminary remarks in mind I am now ready to outline how one could give a comprehensive account of intensionality in language. Consider first the apparent violation of Leibniz's law produced by co-denoting names (or indexicals). If names (indexicals) have no description content, then, as I will argue in §39, these apparent violations of Leibniz's law are pragmatic, not semantic, phenomena akin to those responsible for Mates' puzzle (see §18), and they can be given pragmatic explanations in terms of Grice's theory of conversational implicature (see details, §39). On the other hand, if names (indexicals) do have descriptive content, these apparent violations of Leibniz's law are special cases of the apparent violations produced by definite descriptions. However, the latter can be explained away by means of Russell's theory of descriptions.[7] According to this account, the apparent intensionality of a given occurrence of a definite description (e.g., the occurrence of 'the author of *Waverley*' in 'George IV wished to know whether Scott was the author of *Waverley*') is blamed on the apparent intensionality of occurrences of constituent descriptive predicates or formulas (e.g., 'x is an author of *Waverley*'). Thus, each apparent instance of intensionality produced by a singular term either is explained away pragmatically or is reduced to the apparent intensionality of certain occurrences of predicates of formulas. However, according to the proposed extensional analysis of intensional abstraction, every time we seem to have an occurrence of a predicate or formula that violates the principle of substitutivity of equivalents, the guilty occurrences are not genuine occurrences of predicates or formulas at all, and so they do not constitute genuine violations of this substitutivity principle. Given this extensional analysis of intensional abstraction, the substitutivity principle that is relevant to these linguistic contexts is again Leibniz's law. For according to the analysis, each problematic occurrence of a predicate or a formula is actually an occurrence of an expression correlated via the transformation to an occurrence of a singular term. Now, according to the analysis, this occurrence of a singular term denotes the meaning of the predicate or formula to which it is correlated syntactically. Consequently, the apparent intensional occurrence of a predicate or a formula could lead to a genuine violation of extensional logic only if there were genuine violations of the principle of the substitutivity of synonymous predicates and formulas, i.e., genuine substitutivity failures for predicates or formulas that have the same meaning. However, violations of this substitutivity principle would be instances either of the paradox of analysis or of Mates' puzzle and, hence, could be handled by means developed elsewhere (the paradox of analysis, §20; Mates' puzzle, §39) independently of the present issue concerning the thesis of extensionality.[8] And so by this chain of analyses one could eliminate all apparent violations of the principles of extensional logic. Logic and language would be at bottom extensional. Apparent intensional language would be extensional language concerning intensional entities.

In the remainder of this section I will spell out the details of this extensional analysis of intensional abstraction. In the next section I will take up the topic of meaning, which has figured informally throughout our discussion of the thesis of extensionality and indeed throughout all the preceding chapters. In the final section of this chapter I will, as promised, take up names, indexicals, and Mates' puzzle.

I begin by constructing a first-order extensional language L.*

* Some readers might wish to skip over this technical material.

The primitive symbols of L are the following:

Operators:	$\&, \neg, \exists$
Predicates:	$\mathscr{Conj}^3, \mathscr{Neg}^2, \mathscr{Exist}^2, \mathscr{Exp}^2,$ $\mathscr{Inv}^2, \mathscr{Conv}^2, \mathscr{Ref}^2, \mathscr{Pred}^3$ $F_1^1, F_2^1, F_3^1, \ldots, F_p^q$
Names:[9]	$\bar{F}_1^1, \bar{F}_2^1, \bar{F}_3^1, \ldots, \bar{F}_p^q$
Variables:	$x, y, z, \ldots$
Punctuation:	(,).

The formulas of L are built up in the usual way. As in L_ω, the predicate F_1^2 is singled out as a distinguished logical predicate and is to be rewritten as $=$. The predicates $\mathscr{Conj}^3$, $\mathscr{Neg}^2$, $\mathscr{Exist}^2$, $\mathscr{Exp}^2$, $\mathscr{Inv}^2$, $\mathscr{Conv}^2$, $\mathscr{Ref}^2$, $\mathscr{Pred}^3$ are also singled out as distinguished logical predicates. The semantics for L is done in the usual Tarskian extensional manner with certain added conditions (set forth in a moment) which insure the proper interpretation of these additional logical predicates. Since these added conditions only narrow the class of models for L, every L-formula that is true in all Tarskian models will be true in all models in the narrower class. Thus, any L-formula provable in standard first-order quantifier logic with identity will be valid on the intended semantics for L; this is true in particular for all instances of Leibniz's law and the substitutivity of equivalent formulas. In this precise sense, then, L is a fully extensional language. Notice also that L satisfies Davidson's finite learnability requirement since it has a finite number of primitive constants. The narrowed class of interpretations of L may be obtained in a short-cut way. Consider an arbitrary Tarskian model $\langle \mathscr{D}, \mathscr{R} \rangle$ for L. ($\mathscr{D}$ is the universe of discourse, and $\mathscr{R}$ is an extensional interpretation of the primitive predicates of L relative to $\mathscr{D}$.) Then, the model $\langle \mathscr{D}, \mathscr{R} \rangle$ is *admissible* if and only if it is associated with a standard model $\langle \mathscr{M}, \mathscr{I} \rangle$ for L_ω such that the following rules hold:[10]

$$
\begin{array}{lll}
\mathscr{D} = \mathscr{D}_{\mathscr{M}} & \mathscr{R}(\mathscr{Conj}^3) = \text{Conj} & \mathscr{R}(\mathscr{Inv}^2) = \text{Inv} \\
\mathscr{R}(F_i^j) = \mathscr{G}(\mathscr{I}(F_i^j)) & \mathscr{R}(\mathscr{Neg}^2) = \text{Neg} & \mathscr{R}(\mathscr{Conv}^2) = \text{Conv} \\
\mathscr{R}(\bar{F}_i^j) = \mathscr{I}(F_i^j) & \mathscr{R}(\mathscr{Exist}^2) = \text{Exist} & \mathscr{R}(\mathscr{Ref}^2) = \text{Ref} \\
 & \mathscr{R}(\mathscr{Exp}^2) = \text{Exp} & \mathscr{R}(\mathscr{Pred}^3) = \bigcup_{k \geq 0} \text{Pred}_k .
\end{array}
$$

An admissible model $\langle \mathscr{D}, \mathscr{R} \rangle$ for L is type 1 (type 2) if and only if it is associated in this way with an L_ω-model $\langle \mathscr{M}, \mathscr{I} \rangle$ in which the model structure $\mathscr{M}$ is type 1 (type 2). Relative to the admissible models for L one may give a standard Tarskian definition of truth for L. Finally, relative to the notion of truth for L, the appropriate notions of validity may then be defined: a formula of L is valid$_1$ (valid$_2$) if and only if it is true in all admissible type 1 (type 2) models for L.

Using L's distinguished logical predicates we may contextually define certain functional constants 'conj', 'neg', 'exist', 'exp', 'inv', 'conv', 'ref', 'pred$_0$', 'pred$_1$', ..., 'pred$_k$', ... (The fact that $\mathscr{R}(\mathscr{P}red^3)$ partitions into the predication operations Pred$_0$, Pred$_1$, ... is what makes the functional constants 'pred$_k$' definable in L: pred$_k(x, y) = z$ *iff*$_{\text{df}}$ $\mathscr{P}red^3(x, y, z)$ and z's degree is greater than the degree of x by $k-1$.) As long as L is interpreted in the admissible ways, these contextual definitions will always pick out the intended fundamental logical operations Conj, Neg, ... Now let $[A]_\alpha$ be any intensional abstract of L$_\omega$. Recall the inductive definition of the denotation function given in §14. Let the definition of $D_{\mathscr{I}\mathscr{A}\mathscr{M}}$('$[A]_\alpha$') be written out fully so that no occurrences of $D_{\mathscr{I}\mathscr{A}\mathscr{M}}$ remain in the definiens. The resulting expression θ consists of (1) quotation names of predicates and variables of L$_\omega$, (2) 'Conj', 'Neg', 'Exist', 'Exp', 'Inv', 'Conv', 'Ref', 'Pred$_0$', 'Pred$_1$', ..., and (3) '$\mathscr{I}$', '$\mathscr{A}$', commas, and parentheses. Make the following changes in θ: (1) replace each quotation name of the predicate F_i^j with the associated name $\bar{F}_i^j$ and each quotation name of a variable with the variable itself; (2) replace 'Conj', 'Neg', ... with the associated functional constant 'conj', 'neg', ... which was contextually defined in L; (3) delete '$\mathscr{I}$' and '$\mathscr{A}$' and all associated occurrences of parentheses. The resulting complex expression θ^* is one of the function-*cum*-argument terms that were defined contextually in L. The L$_\omega$-term $[A]_\alpha$ is then defined in L as follows: $[A]_\alpha =_{\text{df}} \theta^*$.

Some elementary examples should help to illustrate how this definition works. Consider the L$_\omega$-term $[F_1^1x]^x$. The result of expanding the definition of $D_{\mathscr{I}\mathscr{A}\mathscr{M}}$('$[F_1^1x]^x$') is Pred$_0(\mathscr{I}('F_1^1'), \mathscr{A}('x'))$. Then after the steps (1)–(3) above, the result is pred$_0(\bar{F}_1^1, x)$. So, one gets $[F_1^1x]^x =_{\text{df}}$ pred$_0(\bar{F}_1^1, x)$, i.e., the proposition that F_1^1x $=_{\text{df}}$ the proposition that is the absolute predication of F_1^1-ness of x. For a second example, consider the L$_\omega$-term $[(\exists x)(F_1^1x \;\&\; F_2^1x)]$. The definition of $D_{\mathscr{I}\mathscr{A}\mathscr{M}}$('$[(\exists x)(F_1^1x \;\&\; F_2^1x)]$') is Exist(Conj($\mathscr{I}('F_1^1')$, $\mathscr{I}('F_2^1')$)), which, after application of the steps (1)–(3), yields exist(conj($\bar{F}_1^1$, $\bar{F}_2^1$)). So, $[(\exists x)(F_1^1x \;\&\; F_2^1x)] =_{\text{df}}$ exist(conj($\bar{F}_1^1$, $\bar{F}_2^1$)), i.e., the proposition that something is both F_1^1 and $F_2^1 =_{\mathrm{df}}$ the proposition that is the existential generalization of the conjunction of F_1^1-ness and F_2^1-ness.

Given the intended semantics for L, the adequacy of the proposed extensional definition of intensional abstraction is confirmed by the following little theorem, which has a straightforward inductive proof:

> Let $\langle \mathscr{M}, \mathscr{I} \rangle$ be an L_ω-model; $\langle \mathscr{D}, \mathscr{R} \rangle$, the associated admissible L-model; $\mathscr{A}$, an assignment; $[A]_\alpha$, an L_ω-term, and θ^*, the translation of $[A]_\alpha$ into L. Then, $[A]_\alpha$ denotes, relative to $\langle \mathscr{M}, \mathscr{I} \rangle$ and $\mathscr{A}$, the same entity as θ^* denotes, relative to $\langle \mathscr{D}, \mathscr{R} \rangle$ and $\mathscr{A}$.[11]

What then of the truth of the thesis of extensionality? This thesis is of course a philosophical thesis, as is, for example, Church's thesis on effective computability. As such, it cannot be proven or disproven. But the prospect of its truth has looked dim to most people recently since even its technical feasibility has seemed out of reach. What I have just shown is that there are no technical barriers to its truth. Its truth turns instead on an ongoing methodological conflict in logical theory.

In closing, I should like to emphasize that the intensional ontology of PRPs is, ironically, what makes possible the defense of the thesis of extensionality. The moral is that those who wish to be extensionalists in logic may be so, but only if they are intensionalists in ontology. Short of artificial limitations on the natural domain of logic, this conclusion seems unavoidable.

38. Semantics

Semantics is the theory concerning the fundamental relations between words and things. Up to now I have been making free and uncritical use of a style of semantics that was developed by Tarski. In Tarskian semantics one defines what it takes for a sentence in a language to be true relative to a model. This puts one in a position to define what it takes for a sentence in a language to be valid. Since validity is what interests logicians, Tarskian semantics often proves quite useful in logic. Despite this, Tarskian semantics neglects meaning, as if truth in language were autonomous. This seems wrong, and it is time to address this issue.

Under the leadership of Donald Davidson,[12] many philosophers and linguists propose to identify the theory of meaning with the theory of truth—or, alternatively, to eliminate entirely the theory of meaning in favor of the theory of truth. The central problem with this approach is that it seems unable to characterize basic facts about what words mean. For example, it seems unable to give a satisfactory explanation of how '$(\forall x)(Ax \equiv Bx)$' can be both true and different in meaning from '$(\forall x)(Ax \equiv Ax)$'. For '$(\forall x)(Ax \equiv Bx)$', if true, has the same truth conditions as '$(\forall x)(Ax \equiv Ax)$'. Learning the meaning of '$(\forall x)(Ax \equiv Bx)$'—or learning how to use '$(\forall x)(Ax \equiv Bx)$'—requires more than learning that this sentence is true if and only if $(\forall x)(Ax \equiv Ax)$.[13]

According to commonsense semantics, the theory of truth is not autonomous from the theory of meaning, unlike what Tarskian semantics would suggest. Instead, the theory of truth is a derived theory obtained from the theory of meaning:

'A' is a true sentence iff_{df} 'A' expresses a true proposition.

'A' is true if and only if what 'A' means is true. (What it takes for a proposition to be true is an antecedent question to be settled by the theory of PRPs; see §45.) Thus, once commonsense semantics is properly developed, the Tarskian theory of truth, though useful in mathematical logic, becomes inessential to the semantics for natural language.

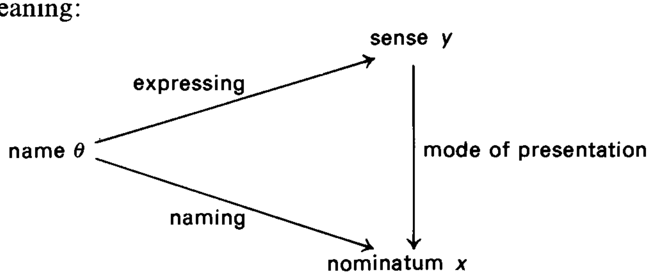
Commonsense semantics takes *meaning*, *naming*, and *expressing* to be the fundamental relations between words and things. (Referring is another relation between words and things that commonsense semantics deems important. I will take up the relation of referring later in this section.) True enough, commonsense semantics has never been given an adequate rigorous formulation. Aside from this, though, none of its critics has been able to make good his claim that the basic commonsense concepts of meaning, naming, and expressing are unsuited to their charge.[14] Because of this, and because the commonsense theory is simpler and more elegant than its competitors, it is the theory one ought to try to develop.

One should first get clear about how meaning, naming, and expressing are related to each other. Since other semantic theories may be viewed as variations on the two classical theories—Frege's and Russell's[15]—my discussion will center on the two classical theories. I will argue that Frege's theory has several faults which are avoided by Russell's theory and that Frege's theory has no advantages over Russell's.

According to Frege, every primitive symbol that is not a variable, and every well-formed string of symbols that does not contain free variables, is a name. Furthermore, according to Frege, there exist *two* fundamentally different meaning relations: expressing and naming. Every name expresses something. This kind of meaning is called a *sense* (*Sinn*). And, at least in an ideal language, every name names something. This kind of meaning is called a *nominatum* (*Bedeutung*).[16] The relations of expressing and naming satisfy the following further general principle:

> θ names x if and only if there is a y such that θ expresses y and y is a mode of presentation of x.

Here Frege takes *mode of presentation* (*Art des Gegebenseins*) to be an unanalysed extralinguistic relation akin to the relation of representation posited by representationalists. (See §42 for more on representationalism.) Thus, Frege gives us the following picture of meaning:



Frege's theory is beset with many difficulties. To begin, it offends our semantic common sense. Verbs such as 'repeats', 'runs', 'chews': is it credible that these name (or refer to) something, as Frege's theory requires? Is it any more credible that sentences, like 'The cat is on the mat', name (or refer to) something? Common sense grants that verbs and sentences do *express* something; but they have no other kind of meaning than this.[17] It is also common sense that names name something (unless they are vacuous). Frege's theory holds that, in addition, a name expresses something that determines what the name names. Although intuition is less clear here, there are provoking arguments (such as those in Kripke's

'Naming and Necessity') that names express no such thing. These offences to common sense only begin the difficulties for Frege's theory. Others arise when the theory is used in his construction of a semantics for *prima facie* intensional languages. Frege's infinite sense-hierarchy renders the resulting syntax forbiddingly complex. And the theory runs into perhaps insuperable difficulties in connection with quantifying-in and with Davidson's finite learnability requirement.

This is where Russell's theory of meaning comes in, for it is in the clear on the above counts. According to the theory, there is only one meaning relation, one fundamental relation between words and things. There is, however, a fundamental *syntactic* distinction between names, on the one hand, and verbs and sentences, on the other. This syntactic distinction is then used to define the subsidiary semantic relations of naming and expressing. An expression θ names x if and only if θ is a name that means x. ("When you said 'John', whom did you mean?") And an expression θ expresses x if and only if θ is a verb or an open or closed sentence that means x. Naming is simply meaning restricted to names, and expressing is simply meaning restricted to verbs and sentences. So Russell gives us the following simple picture:

$$\text{name, verb, or sentence } \theta \xrightarrow{\text{meaning}} x.$$

Whence, if one wishes, one can derive:

$$\text{name } \theta \xrightarrow{\text{naming}} x$$

$$\text{verb or sentence } \theta \xrightarrow{\text{expressing}} x.$$

At the same time, Russell's theory provides a semantics for intensional language. Whether one treats intensional abstraction as undefined (as in the intensional language L_ω) or as defined (as in the extensional language L of §37), the semantics will be a standard Russellian semantics. Either way, the approach runs into no difficulties over quantifying-in or over Davidson's finite learnability requirement.

So far so good. But is Russell's theory as adequate as Frege's? After all, Frege was not led to his theory only by architechtonic concerns. There are two arguments standardly given in behalf of

Frege's theory of meaning. The first is that the theory is an essential ingredient in the extensional analysis of intensional language. However, I have already given just such an analysis without Frege's infinite sense-hierarchy, so the first argument fails. The second argument is that only Frege's theory is able to explain satisfactorily certain elementary meaning phenomena epitomized by Frege's famous '$a = a$'/'$a = b$' puzzle, to which I now turn.

In 'Funktion und Begriff' and 'Über Sinn und Bedeutung' Frege poses a question that may be put as follows: how can '$a = b$' be true yet different in meaning from '$a = a$'? Frege believes that an adequate answer to this question requires invoking his two-kinds-of-meaning semantics. According to this semantics, although sense determines nominatum, nominatum does not determine sense. Thus 'a' and 'b' can differ in sense even when they have the same nominatum. For this reason, '$a = a$' and '$a = b$' also can differ in sense even when they have the same nominatum (i.e., truth value).

Russell believes that Frege's question can be adequately answered without positing two kinds of meaning. Sticking to his own one-kind-of-meaning semantics, Russell answers Frege's question by means of a two-part *syntactic* theory. First, he holds that, if '$a = b$' is true but different in meaning from '$a = a$', then 'a' or 'b' is an overt or covert definite description or extensional abstract. Secondly, Russell holds that definite descriptions and extensional abstracts are incomplete symbols[18]—definite descriptions being analysed with the theory from 'On Denoting' and extensional abstracts, with the theory from *Principia Mathematica*.[19] The effect of this procedure is to shift the weight of the explanation away from expressions that would seem to be subjects; Russell instead places the weight of the explanation on underlying predicates and formulas. The strategy succeeds because the only kind of meaning that a predicate or a formula can in any natural sense be said to have is the intensional entity that it expresses. The fact that such intensional entities can be equivalent without being identical is what makes it possible for the predicates and formulas to be equivalent without being identical in meaning. This, in turn, is what makes it possible for the initial '$a = a$'/'$a = b$' sentences to be true without being identical in meaning. (In what follows I will focus on definite descriptions since, if Russell's answer works for them, then it also works for extensional abstracts.)

Over the years there has been a wide range of doubts about the adequacy of Russell's answer to Frege's question. Let me begin by considering the doubts aimed at the second part of Russell's theory. These doubts fall into two main kinds.

First, they arise in connection with the paradox of analysis. (For example, if it is possible for someone to believe that G(the F) and yet not believe that $(\exists v)((Fu \equiv_u v = u) \,\&\, Gv)$, would this show that Russell's analysis of definite descriptions is in error?) These doubts, however, arise for analyses in general and are not a special problem for Russell's analysis of definite descriptions. Any adequate resolution of the paradox of analysis should handle doubts about the special case of Russell's analysis. These doubts can be allayed, for example, by an adaptation of the resolution offered in §20.

Secondly, there are doubts, prompted by Strawson's 'On Referring' and Donnellan's 'Reference and Definite Descriptions', concerning Russell's view that definite descriptions are semantically incomplete symbols. Yet on this matter there is a forceful defense of Russell based on the methods developed by Paul Grice in 'Logic and Conversation' and 'Definite Descriptions in Russell and in the Vernacular'. On the picture that emerges, although definite descriptions usually do have a reference, referring, unlike naming, is a *pragmatic* relation not a *semantic* relation. Thus, definite descriptions, while pragmatically complete symbols (they typically refer in conversational contexts), are semantically incomplete: their being co-referential in conversational context does not make them alike in any kind of genuine semantic meaning.[20]

I now come to the doubts aimed at the first part of Russell's answer to Frege's question, i.e., Russell's theory that, if '$a = b$' is true yet different in meaning from '$a = a$', then 'a' or 'b' is an overt or covert definite description (or extensional abstract). Two main doubts arise here. The first springs from the Mill-Kripke doctrine that ordinary names have no descriptive content (see Kripke, 'Naming and Necessity'). But Fregean semantics, not Russellian semantics, is this doubt's proper target. Indeed, the Mill-Kripke doctrine is a straightforward consequence of Russellian semantics when that theory is conjoined with the syntactic theory that ordinary names are genuine names. (An analogous doubt concerns indexicals. Throughout this section I omit discussion of indexicals since what I say about Russell on ordinary names applies analogously to Russell on indexicals. Indexicals will be explicitly discussed in the following section.)

Finally, there is the doubt raised by Alonzo Church,[21] the most prominent American advocate of Frege's theory. The thrust of Church's doubt seems to be this. Let it be granted that any isolated instance of Frege's puzzle can be explained in a Russellian framework by bringing to the surface occurrences of covert definite descriptions. Nonetheless, such a procedure must use new constants, among which are often new predicates. Nothing in Russell's theory guarantees that such new constants themselves will not be responsible for further instances of Frege's puzzle. Now although such further instances of Frege's puzzle might, in turn, be explained by still further use of Russell's methods, this process obviously should not go on forever. But nothing in Russell's theory insures that it will not. Frege's approach, in contrast, runs into no comparable difficulty.

Let me take a moment to show that Church's doubt is unfounded. The point of Russell's theory (i.e., that overt or covert definite descriptions are responsible for all instances of Frege's puzzle) is that it permits him to shift the weight of the explanation onto the predicates and formulas embedded in those descriptions. Church believes that in higher-order languages (such as those fashioned after the language of *Principia Mathematica*) these predicates and formulas could generate new instances of Frege's puzzle. The reason is that in higher-order languages predicates and formulas are linguistic subjects (see §23), making strings such as '$F = F$', '$F = G$', '$A = A$', and '$A = B$' well-formed. It does not follow from this, of course, that Church's doubt applies to first-order languages with Russellian semantics. Though Russell's procedure shifts the weight onto predicates and formulas, this can generate no such instances of Frege's puzzle since predicates and formulas are just not linguistic subjects in a first-order setting. So Church's doubt that Russell's theory can handle all instances of Frege's '$a = b$'/'$a = a$' puzzle all but evaporates when one properly distinguishes Russell's theory of meaning from the combined theory consisting of Russell's theory of meaning and the *Principia Mathematica* theory that logical syntax is at bottom higher-order. Indeed, I argued (§§10, 22–6) that the higher-order approach is defective on several counts; logical syntax is at bottom first-order.

All that could now keep Church's doubt alive is the worry that there are other kinds of *prima facie* linguistic subjects (besides spurious higher-order "names") that render Russell's theory of meaning less adequate than Frege's. But what could these conceivably be? Concerning intensional abstracts,[22] extensional abstracts, definite descriptions, and ordinary functional constants,[23] it is clear that the Russellian theory is at least as adequate as Frege's. So the question comes down to the issue of ordinary names (and indexicals). But it is easy to show that in its treatment of these expressions, Russell's theory is again at least as adequate. There are two relevant theories on the content of ordinary names, Frege's and Mill's. According to Frege's theory, associated with each ordinary name is a descriptive content that serves to determine its nominatum; according to Mill's theory, ordinary names lack such a content. I will consider each of these theories in turn.

Suppose, with Frege, that each ordinary name '*a*' has an associated descriptive content. Then, according to Russell's syntactic theory, '*a*' is not a genuine name; genuine names do not express anything. So according to Russell, '*a*' should be treated as a disguised definite description: $a =_{\mathrm{df}} (\iota v_i)F^1_k(v_i)$, where the predicate F^1_k is interpreted so as to express the property Fregeans would associate with '*a*'. (Conventions governing scope and the introduction of the new variable v_i are naturally in force.) This familiar maneuver enables the Russellian to handle ordinary names if they do have associated descriptive content. On the other hand, suppose, with Mill, that they do not. In this case, ordinary names may be simply treated as special undefined singular terms rather akin to variables with fixed assignments. Now since on this theory ordinary names '*a*' and '*b*' have no descriptive content, it is not possible for '$a = b$' to be true yet different in semantic meaning from '$a = a$'. So no genuine Fregean puzzles could arise.[24] Thus, whether or not ordinary names have associated descriptive contents, a Russellian semantics for them is as adequate as a Fregean semantics once the proper first-order logical syntax is identified. Russell's theory of meaning, therefore, is at least as adequate as Frege's. This shows that Church's doubt about Russell's theory of meaning is unfounded.

Since the Russellian account of the '$a = a$'/'$a = b$' puzzles is at least as adequate as Frege's and since Russell's theory of meaning is superior to Frege's theory on the several other counts reviewed above, Russell's is the better theory. If Frege's theory on the descriptive content of names is right, then the Russellian will treat all names as abbreviations for contextually defined definite descriptions. If one does this in a first-order extensional setting, the resulting language will contain no names. Using the algebraic apparatus of §§13–14, one can define the Russellian meaning function $\mathrm{M}_{\mathscr{I}\mathscr{M}}$ for this language simply as follows:

$$\mathrm{M}_{\mathscr{I}\mathscr{M}}(F_i^j) = \mathscr{I}(F_i^j)$$
$$\mathrm{M}_{\mathscr{I}\mathscr{M}}(A) = \mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}([A]_{v_1 \ldots v_j})$$

where $\mathscr{M}$ is an algebraic model structure, $\mathscr{I}$ is an interpretation, $\mathscr{A}$ is an assignment, and $v_1, \ldots, v_j$ are the free variables of A in order of their first free occurrences.[25] One may then use $\mathrm{M}_{\mathscr{I}\mathscr{M}}$ to define the Fregean sense function $\mathrm{S}_{\mathscr{I}\mathscr{M}}$ and the Fregean reference function $\mathrm{R}_{\mathscr{I}\mathscr{M}}$:

$$\mathrm{S}_{\mathscr{I}\mathscr{M}}(\theta) = \mathrm{M}_{\mathscr{I}\mathscr{M}}(\theta)$$
$$\mathrm{R}_{\mathscr{I}\mathscr{M}}(\theta) = \mathscr{G}(\mathrm{M}_{\mathscr{I}\mathscr{M}}(\theta))$$

where θ is any predicate or formula.[26] If, on the other hand, Mill's theory of names is correct, then one is to use an analogous procedure except that names are given a meaning in much the same way that one might give variables fixed assignments. In either case, the definability of the Fregean sense and reference functions in terms of the Russellian meaning function shows that Fregean semantics provides no semantic information not already provided by the essentially simpler Russellian semantics.

What makes this Russellian semantics viable is the intensional ontology of PRPs. It is natural to wonder, then, which of the two traditional conceptions of PRPs pertains to the semantics for natural language. Are the meanings of natural language formulas PRPs of conception 1 or conception 2? The work of Paul Grice makes this question easy to answer.[27] Grice is able to define what it is for a speaker to mean a given propostion by performing an intentional action. The definition is given in terms of the intentions with which the speaker performs the action, including in particular his intentions to get his hearers to believe the proposition on the basis of a certain preferred inference route. Using the intentionalist analysis as a first step, Grice is then able to analyse how a sentence or a word comes to mean what it does for a community of speakers. That is, Grice is able to analyse how through their intentional activity a community of speakers comes to invoke an abstract semantical relation (such as the relation $\mathrm{M}_{\mathscr{I}\mathscr{M}}$ that I just characterized) as the meaning relation for the language they actually speak. In order for a given pure abstract semantics to become the semantics for an actual natural language, the speakers of that language must stand in certain complex intentional relations to that pure abstract semantics and, thence, to the meanings isolated by it. Intentionality therefore is the link-up between pure abstract semantics and the semantics for a spoken natural language. Now given this intentionalist analysis, the type of propositions that come to be the meanings of sentences in natural languages must of necessity be the type that are typically intended, believed, remembered, etc. And the type of propositions that typically serve this function are precisely conception 2 propositions. Therefore, it is conception 2 PRPs that pertain to the semantics for natural language.

When I asserted at the beginning of this section that commonsense semantics had never been given an adequate rigorous formulation, I had two things in mind. First, previous attempts to rigorously formalize commonsense semantics have mistakenly formalized Frege's semantics, not Russell's. Yet Russell's is the commonsense theory. Secondly, previous attempts to formalize commonsense semantics have utilized the possible-worlds technique, which is based on conception 1, not conception 2. Yet conception 2 is the one relevant to the commonsense semantics for natural language. These deficiencies have now been remedied.

39. Pragmatics

I now shift to four problems in the logic for intentional matters: (1) *prima facie* substitutivity failures involving co-denoting names, (2) *prima facie* substitutivity failures involving co-denoting indexicals, (3) Mates' puzzle (which concerns *prima facie* substitutivity failures involving synonymous predicates and formulas),[28] and (4) Geach's problem of intentional identity (which concerns *prima facie* quantification over non-actual possibilia).[29] Recently, there has been much provocative investigation of these issues;[30] to attempt definitive solutions of them at this point would be premature. I aim only to explore some candidate solutions. My purpose in doing so is to convey the explanatory power of the theory of PRPs and, in turn, to suggest that this theory provides a general framework within which these four problems can be solved eventually. If I am right, solving these problems is best viewed as a matter of fine-tuning among the applications of the theory, fine-tuning that does not threaten the theory's underlying Platonistic character.

Let us consider names and indexicals first. I have shown that Russell's theory of meaning is compatible with both classical theories on the content of ordinary names (and indexicals), Frege's theory and Mill's theory. It is not the job of the theory of PRPs to decide the Frege/Mill controversy; all that matters is that the theory be adaptable to whichever doctrine is correct. Assuring oneself of its adaptability is relatively straightforward if Frege's doctrine is correct; for then the substitutivity problems involving such singular terms either submit to traditional solutions (Frege's or Russell's) or collapse into instances of Mates' puzzle, which can be dealt with in the manner suggested later in this section. If, on the other hand, Mill's theory is right, then the substitutivity problems involving names (and indexicals) would call for a non-traditional approach. In a moment I will propose such an approach. If it is successful, then the adequacy of the framework of PRPs is guaranteed, no matter which theory on the content of ordinary names (and indexicals) is correct.

Suppose for the sake of discussion that Mill's theory is correct. Let us compare our three substitutivity problems—Mates' puzzle and the two involving primitive singular terms. Recall from §18 that, unlike the paradox of analysis, whose source is ignorance of definitions of concepts, Mates' puzzle originates in some form of linguistic (or historical or social) ignorance. For example, consider someone x who knows what the verb 'chew' expresses but is ignorant that the verb 'masticates' expresses the same thing. In this situation it might be natural to affirm

(1) x does not know that whatever masticates chews.

while denying

(2) x does not know that whatever chews chews.

And this is so despite the fact that the literal Russellian meanings of sentences (1) and (2) are the same.[31] Now according to Mill's theory, ordinary names have no descriptive content; they only name. If this is right, observe how similar Mates' problem is to the substitutivity problem for co-denoting ordinary common names in intentional sentences. The types of ignorance responsible for these problems appear analogous. E.g., ignorance that whatever masticates chews would seem quite on a par with ignorance that pot is marijuana, that consumption is tuberculosis, that lorries are trucks, that pumas are cougars, that filbert is hazelnut, etc. And if this is so for the substitutivity problem for ordinary common names ('pot', 'consumption', 'lorry', 'puma', etc.), what reason could there be for thinking that the substitutivity problem for ordinary proper names is different? It would seem that one has no choice but to treat ignorance that Scott is Sir Walter, that Tully is Cicero, etc. on a par with ignorance that whatever masticates chews. (Of course, one must be careful to distinguish genuine names from descriptions masquerading as names. For example, 'H_2O' is no name but a description short for something like 'the compound whose molecules bind together 2 hydrogen and 1 oxygen atom'.[32]) In fact, if both Mill's doctrine and Russellian semantics are right, one can all but prove that the substitutivity problem for all ordinary names—proper as well as common—is on a par with Mates' substitutivity puzzle. The argument goes as follows. The only difference between the two problems lies in the fact that the former problem involves ordinary names whereas the latter problem involves verbs. However, given Russell's theory of meaning, there is only one kind of meaning, and genuine names stand in this same meaning relation to their meanings as do verbs to theirs. Further, according to Mill's theory, ordinary common and proper names are genuine names. It follows that ordinary names stand in the very same meaning relation to their meanings as do verbs to theirs. Hence, there can be no relevant *semantical* difference between ordinary names and verbs. Consequently, there can be no relevant semantical difference between the substitutivity problem for ordinary names and Mates' puzzle. The only difference between the two problems therefore is *syntactical*: one concerns ordinary names and the other concerns verbs. Indeed, the two problems are really species of the same general problem: the problem of the substitutivity of synonyms.

Much the same thing holds for the substitutivity problem involving demonstratives. For example, suppose that a given speaker utters 'this' while pointing directly to a certain object in plain view and 'that' while pointing through some complex optical apparatus to what turns out to be the same object. In this situation it might be natural for this speaker to deny

(3) I believe that this = that.

while affirming

(4) I believe that this = this.

Yet from a semantical point of view demonstratives have elusive descriptive content. If none can be found, it would seem that, like the substitutivity problem for Millian names, the substitutivity problem for demonstratives is just a syntactical variant of Mates' puzzle and, hence, that all three substitutivity problems are species of the same general problem. If this is so, the three problems call for a unified solution.

I will now sketch a metaphysical theory of belief which should bring one closer to a solution. (This theory easily generalizes to the other problematic intentional relations.) Consider a normal conversational context in which it would be appropriate to utter 'x believes that A'. Typically, the believer x must satisfy *two* conditions. First, he must be what I call *cognitively committed* to the proposition that is the meaning of the embedded sentence A; he need not be acquainted (in the traditional epistemological sense[33]) with this proposition, however. Secondly, x must be convinced of a proposition (often, but not always, a different one) with which he is acquainted. (I call this a *conviction in acquaintance* or *conviction$_{acq}$* for short.) It is in virtue of this conviction in acquaintance that x is cognitively committed to the proposition literally expressed by A. Now a person is cognitively committed to all those propositions of which he is convinced in acquaintance, but the converse does not hold. This is crucial. In daily social intercourse, when we rely on another believer for information about the world, we focus on those of his cognitive commitments of which he is not convinced in acquaintance. For these cognitive commitments deal directly with the objects in the world and ignore the individual modes of epistemic access to those objects, which usually are of no special interest and which also are difficult for us to discover. However, when we wish to explain a believer's actions, we ultimately look to his convictions in acquaintance. For these are what figure in his deliberations about what to do; he typically is not even immediately aware of many of his associated cognitive commitments.[34]

Let us apply this metaphysical scheme to some problematic cases. For example, suppose that I have severe amnesia and that I sincerely utter the sentence 'I believe that I am not George Bealer'. In this situation I would be cognitively committed to the proposition literally expressed in the context by the embedded sentence 'I am not George Bealer'; i.e., I would be cognitively committed to the necessarily false proposition $[x \neq x]^x$, where I am x. But I would be convinced in acquaintance of a different proposition, perhaps $[x \neq$ "George Bealer", as he is called$]^x$, where I am x. Or even better, simply $[x \neq$ "George Bealer"$]^x$, where the quotation marks are "scare quotes". (Without taking a position on the proper analysis of scare quotes, one may be confident that it involves some form of metalinguistic allusion.) It is in virtue of such a conviction$_{\text{acq}}$ that I would be cognitively committed to the necessary falsehood $[x \neq x]^x$. Of course, I would not be convinced$_{\text{acq}}$ of $[x \neq x]^x$; that would take gross irrationality whereas I only suffer from amnesia.[35]

Or consider the example given in Kripke's 'A Puzzle About Belief' (pp. 254 ff). A Frenchman Pierre, who has only seen photos of London, sincerely utters the sentence 'Londres est jolie'. Later Pierre moves to London. After learning English, he sincerely utters the sentence 'London is not pretty', but he does so without knowing that 'Londres' and 'London' name the same city. Indeed, he still would sincerely utter 'Londres est jolie'. Using the above metaphysical scheme one would say that Pierre is cognitively committed to both the proposition that London is pretty and its negation, i.e., to both $[u \, \Delta \, v]^{uv}$ and $[u \not\Delta v]^{uv}$, where $u =$ London and $v =$ prettiness. He is cognitively committed to the first proposition in virtue of being convinced in acquaintance of a further proposition—e.g., the proposition that "Londres", as it is called, is pretty. And he is cognitively committed to the negation of the first proposition in virtue of being convinced in acquaintance of still another proposition—e.g., the proposition that "London", as it is called, is not pretty. (As before, I am using scare quotes.) Pierre's logical acumen is not under suspicion, however, for logical acumen gets tested only against those cognitive commitments that a person is immediately aware he has, such as those that are convictions$_{\text{acq}}$. And the two propositions of which Pierre is convinced in acquaintance are logically independent.

Before going further, I should meet a possible worry.[36] Suppose in the above story that Pierre is a rather primitive fellow who has never articulated any of the metalinguistic concepts belonging to linguistic theory. This would show, so the worry supposes, that Pierre could not be convinced in acquaintance of the propositions ["Londres", as it is called, is pretty], ["London", as it is called, is not pretty], or anything like that, for such propositions appeal to metalinguistic concepts. This line of argument, however, overlooks the paradox of analysis. Consider an analogy. Suppose that Pierre has never articulated a geometric theory and that he is ignorant of how to define (analyse) what a circle is. Still, if Pierre is convinced$_{acq}$ that there are circles, then one can truly say ‘Pierre is convinced$_{acq}$ that there are loci of points in the same plane equidistant from a common point’. And this is so even though Pierre might be brought up short by the utterance. Using the apparatus developed for resolving the paradox of analysis (see §20), one can easily explain what is going on. There are *two* different propositions denoted by the ‘that’-clause ‘that there are loci of points . . .’:

$[(\exists y)y \text{ is a locus of points} \ldots]$
$[(\exists y)\underline{y \text{ is a locus of points} \ldots}]$.

Pierre has no conviction$_{acq}$ concerning the former proposition, for it involves the analysed concept of circularity. However, he does have a conviction$_{acq}$ concerning the latter proposition; this proposition (which is just the proposition $[(\exists y)\underline{y \text{ is a circle}}]$) does not involve an articulated definition (analysis) of what a circle is. Now on analogy, the fact that Pierre lacks an articulated linguistic theory provides no evidence whatsoever that he lacks convictions$_{acq}$ concerning metalinguistic propositions. For Pierre might simply be ignorant of how to define (analyse) the metalinguistic propositions with which he is acquainted. (This sort of ignorance is surely pervasive.) Of course, care must always be taken in the formal statement of unanalysed convictions$_{acq}$. In Pierre's case, for example, we might want to represent his convictions$_{acq}$ with something like the following:

$[\underline{\text{“Londres” as it is called}} \text{ is pretty}]$,
$[\underline{\text{“London” as it is called}} \text{ is not pretty}]$.[37]

With this worry allayed, let us now consider how intentional verbs, e.g., ‘believe’, behave in natural language. There are two idealized positions on this question. The first is that ‘believe’ literally expresses a concept that applies only to what one is convinced of in acquaintance.[38] The second is that it expresses a concept that applies only to one's cognitive commitments. Now I am rather persuaded that in everyday speech each of these concepts operates at least pragmatically, if not semantically. Indeed, I would not be surprised if in ordinary language ‘believe’ expresses both concepts, or even some composite of them. At the same time, I am persuaded that with suitable maneuvering either of these two idealized positions can be made to fit all the linguistic data, and that this can be done within the general framework provided by the theory of PRPs.[39] I need not decide here which position is best. For illustrative purposes, though, I will sketch a version of the second position.

On this version 'believes' literally expresses a concept that applies just to what one is cognitively committed to. Yet when a speaker sincerely utters 'x believes that A' in conversation, he typically does two things. First, he asserts that x is cognitively committed to $[A]$—i.e., to the proposition that is the meaning of the sentence A. Secondly, he presupposes that there is some conviction in acquaintance that is the vehicle of x's cognitive commitment to $[A]$. (Such presuppositions arise through mechanisms of the sort isolated by Paul Grice in his conversational pragmatics; see his 'Logic and Conversation' and 'Definite Descriptions in Russell and in the Vernacular'.) In most contexts the identity of this conviction in acquaintance is irrelevant, and the speaker leaves it indefinite. But in some contexts its identity becomes of interest, and the speaker intends to indicate at least roughly what it is. In these contexts the utterance of 'x believes that A' carries a conversational implicature that x has a conviction in acquaintance which falls within the indicated range.

Consider an example to see how this works. If I sincerely utter 'x believes that most pot is grown in Colombia', I assert that x is cognitively committed to the proposition that most pot is grown in Colombia. I also presuppose that x has this cognitive commitment in virtue of some conviction in acquaintance. But just which one, since it is of little importance, I leave indefinite. If I sincerely utter 'x believes that most marijuana is grown in Colombia', my assertion is the same as before, and so is my presupposition; the substitution of 'marijuana' for 'pot' changes neither of these. Next suppose that I sincerely utter 'x believes that pot = pot'. Here, I assert that x is cognitively committed to the proposition that pot = pot, and I presuppose that x has this cognitive commitment in virtue of some conviction in acquaintance. As in the other cases the identity of the conviction$_{acq}$ is of no relevance, so I leave it indefinite. But suppose I sincerely utter 'x believes that pot = marijuana'. Although what I assert about x's cognitive commitment remains unchanged, the conversational pragmatics becomes different. Since 'pot' and 'marijuana' are synonymous, utterances of 'x believes that pot = pot' and 'x believes that pot = marijuana' make the same assertion about x's cognitive commitment. But since the former sentence provides such a simple way to make this assertion, an utterance of the latter sentence signals that there is some special reason for not using the simpler form. Therefore, such an utterance must conversationally implicate something beyond what the sentence expresses semantically. The ripest candidate for this conversational implicature would be something that concerns the conviction in acquaintance underlying x's cognitive commitment. The conversationally salient feature of the sentence is its lexical complexity. This suggests that the implicature concerning x's conviction in acquaintance has something to do with the lexical items themselves. A thing's being called by a certain name is an obvious mode of epistemic access to that thing, so the conversational implicature would often be simply that x's conviction in acquaintance is some metalinguistic proposition such as the proposition that "pot" = "marijuana" (scare quotes again) or something like that. (Of course, depending on the context, non-linguistic modes of epistemic access are often conversationally more relevant than linguistic modes, and the implicature is affected accordingly.) So it is that in this case the substitution of 'marijuana' for 'pot' sharply affects the conversational pragmatics. Utterances of 'x believes that pot = pot' and 'x believes that pot = marijuana' conversationally say quite different things, and this explains why their truth values can differ.

It is no coincidence that in the last case the substitution of 'marijuana' for 'pot' sharply affects conversational pragmatics. True, substitution of co-denoting names—and co-denoting demonstratives and synonymous predicates—usually does not have this sort of pragmatic effect, for in most conversational contexts the interest in an utterance of a belief sentence lies in the literally expressed cognitive commitment. However, in some situations contextual signals shift interest to the conviction in acquaintance underlying the literally expressed cognitive commitment. Here the particular way in which the cognitive commitment is expressed becomes relevant, for it provides contextual cues about the believer's epistemic access to his cognitive commitment and, hence, about what his conviction in acquaintance is. Substitution can thus affect whether an utterance of a belief sentence concerns only the literally expressed cognitive commitment or whether it concerns in addition the underlying conviction in acquaintance, and it can also affect the identity of such conversationally implicated convictions in acquaintance. Substitutions seem invalid in exactly those cases in which they have one of these pragmatic effects, and these are precisely the cases of *prima facie* substitutivity failures I set out in this section to explain. If this is correct, then the three types of substitutivity puzzles—and their explanations—belong to conversational pragmatics, not semantics.[40]

Now the point of the above exercise has not been to give a definitive solution to these problems. Rather, it has been to provide evidence that their solution can be carried out within the general framework provided by the theory of PRPs. I will now attempt a similar exercise for our last puzzle in the logic for intentional matters, namely, Geach's problem of intentional identity.

Geach tells the following little story. A reporter visits a region where there is a rumor that a witch is on the loose. Although the reporter does not himself believe in witches, he makes the following report about the beliefs of two locals:

Hob believes that the witch blighted the sheep, and
Nob believes that she killed the cow.

Geach asks us to assume that there is something true in what the reporter has said; the problem is to characterize what it is. (We need not assume that the uttered sentence is literally true; I doubt that it is.) The problem is one of *intentional* identity because the reporter's statement would seem to imply that Hob and Nob have a belief about the same witch even though no witch exists except, as it were, in the minds of Hob and Nob. But what on earth does this mean?

Someone might try to solve this problem by augmenting the ontology of PRPs with non-actual possibilia, non-existent subsistents, or intentional inexistents. Doing so might permit one to represent the reporter's statement as being about some non-actual, non-existent, or inexistent witch. Though nothing in the theory of PRPs rules it out formally, this strategy taxes the principle of ontological economy, and it does violence to commonsense realism, a view that, if possible, one ought to hold on to.[41]

Where else might one look for a solution to Geach's problem? Evidently the only alternative is to analyse the reporter's statement in such a way that the problematic beliefs of Hob and Nob are propositions that involve descriptive witch-concepts that bear some suitable relation to one another. The simplest example of such a descriptivist analysis is this:

> For some descriptive concept *w*, Hob believes that the *w* witch blighted the sheep and Nob believes that the *w* witch killed the cow.

However, Geach argues against this analysis. His argument is that, though Hob and Nob might conceivably share a descriptive witch-concept, the truth of the reporter's statement does not require that they do.

One might try to meet Geach's argument by formulating a more sophisticated descriptivist analysis. It would be preferable, though, to preserve the form of the initial simple analysis while also doing justice to Geach's intuition that Hob and Nob need not share any descriptive witch-concept. The theory of belief sketched earlier in this section permits this. It is entirely possible that there is no descriptive concept *w* such that Hob is convinced in acquaintance that the *w* witch blighted the sheep and Nob is convinced in acquaintance that the *w* witch killed the cow. This seems to be the basis of Geach's intuition. However, this in no way prohibits Hob and Nob from having other convictions in acquaintance, ones that would give them precisely the sort of cognitive commitments that would validate the simple descriptivist analysis. If Hob and Nob were to have such convictions in acquaintance, Geach's problem about them would be solved.

To make this solution plausible, I will describe a situation in which Hob and Nob would have the relevant sort of convictions in acquaintance and cognitive commitments. First, Hob is convinced in acquaintance that the witch who is at the root of the reference tree to which he is presently a party blighted the sheep. This conviction in acquaintance gives Hob a cognitive commitment to the proposition that the witch who is at the root of reference tree *R* blighted the sheep. Secondly, Nob is convinced in acquaintance that the witch who is at the root of the reference tree to which he is presently a party killed the cow. This gives Nob a cognitive commitment to the proposition that the witch who is at the root of reference tree *S* killed the cow. Finally, reference tree *R* (i.e., the one to which Hob is a party) and reference tree *S* (i.e., the one to which

8

Qualities and Concepts

> By a quality I mean that in virtue of which things are said to be qualified somehow.
>
> Aristotle, *Categories*

40. Qualities, Connections, and Conditions

All objects have countless properties and stand in countless relations. Most of these properties and relations are of little interest, however, for most are not genuine *qualities* or *connections*. Qualities, Aristotle tells us, are that in virtue of which things are said to be qualified. Connections, analogously, are that in virtue of which things are said to be connected. But these remarks taken alone are not very helpful, true though they may be; elucidation is needed.

Examples of properties and relations that are commonly thought to be qualities and connections might be helpful at the outset.[1] There is a widespread belief that certain qualities and connections—called phenomenal qualities and connections—can be known in experience. Examples of such qualities would be colors, tastes, sounds, smells, shapes, textures, hot and cold, and inner feelings (of the sort associated with emotions). And examples of such connections would be the conscious operations of mind themselves, e.g., sensing, feeling, (conscious) thinking, (conscious) wanting, (conscious) deciding, etc. Theories provide another source of examples of properties and relations that are thought to be qualities and connections. In contemporary physics, for example, the quark-theoretic properties of charm, strangeness, color, etc. are typically thought of as qualities. Or in Mendelian genetics traits are thought of as qualities, and the relation of inheritance is thought of as a connection. The relation of gravitational attraction in classical physics, the relation of association in associationist psychology, the relation of stimulation in behavioral psychology, the relations of perceiving, believing, wanting, and deciding in cognitive psychology: these are all thought by the proponents of the respective theories to be connections. One more example of a connection would be the predication relation from logical theory.

Properties and relations that are not genuine qualities and connections may be called *Cambridge* properties and relations.[2] Perhaps the most notorious Cambridge property in recent philosophical literature is the property grue, i.e., the property of being green if examined before t and blue otherwise. An example of a Cambridge relation would be the relation holding between things x and y such that x is green and y is blue.

Examples can help to impart the intuitive distinction between genuine qualities and connections, on the one hand, and Cambridge properties and relations, on the other; but examples only go so far. Something else that can be done is to draw attention to the special roles that qualities and connections play, or at least ought to play, in descriptions of experience and in theories.

It would seem that we experience colors, smells, sounds, hot and cold, inner feelings, the conscious operations of mind, etc. But Cambridge properties we cannot experience; for example, nothing could reasonably count as experiencing grue. In this way phenomenal qualities and connections play a fundamental role in the constitution of experience. Because of this, one's phenomenal descriptions are, or at least should be, given in terms of genuine phenomenal qualities and connections, not Cambridge properties and relations. To dramatize this point, consider an example. Suppose that for a certain duration of time ending at t everything looks green to me and then suddenly everything looks blue. There will have been a distinct change in my experience. This change will be registered in my phenomenal description if that description is given in terms of the qualities green and blue. But if instead the phenomenal description is given in terms of certain Cambridge properties, the change might not be registered. Indeed, the very concept of change in experience would be unintelligible without the logically prior concepts of quality and connection. And much the same thing goes for the concepts of orderliness and disorderliness in experience.

Qualities and connections also play a fundamental role in theories.[3] Changes in the world consist primarily of changes in the qualities and connections of things in the world. So theoretical descriptions and explanations of change, if they are to be adequate, must be given in terms of genuine qualities and connections; Cambridge properties and relations enter in only secondarily.[4] In much the same way, qualities and connections, but not Cambridge properties and relations, play a primary role in the objective, non-arbitrary categorization and identification of objects. Why an object is the particular kind of object it is must be explained in terms of its qualities and connections. And why an object continues to be the same thing that it was earlier must be explained in terms of continuities and changes in its qualities and connections.

The picture that emerges, then, is that qualities and connections are determinants of the phenomenal, causal, and logical order of the world whereas Cambridge properties and relations are idle in these respects.[5]

So far I have tried to give an intuitive indication of what qualities and connections are by providing various candidate examples and by indicating in a rough way the distinctive roles they play, or ought to play, in phenomenal description and in theory. One more way in which I will try to impart the concepts of quality and connection is by indicating the key role they can be expected to have in a solution to Nelson Goodman's new problem of induction.[6]

The degree to which inductive generalizations are epistemologically justified varies widely. Since inductive generalizations are performed on properties and relations, one source of the variability may be traced to the kind of properties and relations involved. For example, inductive generalizations on genuine qualities or connections (e.g., green and blue) have *ceteris paribus* a high degree of justification. And inductive generalizations on Cambridge properties or relations (e.g., grue and bleen) have *ceteris paribus* a low degree.[7] The reason for this is plain. The ideal inductive generalization begins with an observed order and projects it into a general order. But the very concept of orderliness is one that pertains to qualities and connections. When one says that things, observed or unobserved, are orderly, one implies that neat generalizations hold for relevant qualities and connections. After all, qualities and connections are the determinants of the phenomenal, causal, and logical order of the world. Consequently, in formulating the principle of induction one must pay special attention to the properties and relations upon which the inductive generalizations are performed. Goodman's new problem of induction is really just the problem of finding a formulation of the principle of induction that is acceptable in this regard. Even though this problem is not easy, its solution can at least be expected to be straightforward once one has at hand the concepts of quality and connection.

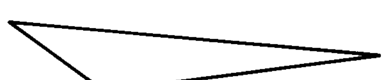
Considerable methodological confusion surrounds Goodman's problem, however. This is generated by a failure to properly distinguish the new problem of induction (i.e., the problem of finding a formulation of the principle of induction that is acceptable regarding the issue just discussed) from two further problems not at all new: one, a traditional metaphysical problem; the other, a traditional epistemological problem. The metaphysical problem is that of giving precise non-circular definitions of the concepts of quality and connection, and the epistemological problem is that of showing how in particular cases to successfully distinguish genuine qualities and connections from Cambridge properties and relations. A few methodological comments on these two traditional problems are in order.

The metaphysical problem of defining the concept of quality goes back at least as far as Aristotle, and the metaphysical problem of defining the concept of connection goes back at least as far as Hume. It is important to understand that these problems do not belong to epistemology (or philosophy of science for that matter), nor will they ever be solved by epistemological means. They fall squarely within traditional metaphysics. Fortunately they are foundational problems there, and this enables their solutions to be given by appealing to logical theory. The tradition of solving foundational metaphysical problems by the means of logical theory was established by Plato and Aristotle and pursued actively in medieval philosophy and has been continued in modern philosophy by such figures as Leibniz, Frege, and Russell. Working in the same tradition, I will employ a logic for PRPs to define the concepts of quality and connection.[8]

Deep traditional roots also underlie the epistemological problem of how in particular cases to distinguish genuine qualities and connections from other properties and relations. This problem—or at least a version of it—can be traced back to Plato, who was concerned with the question of how in particular cases to determine the identity of genuine forms as opposed to spurious ones.[9] Versions of this epistemological problem are also evident in nearly all the classical modern philosophers from Descartes through Kant. Now although this problem is quite difficult, it is not so resistant to solution as relativists think. For given the special role that qualities and connections play in phenomenal description—and in the constitution of experience itself—we may look to our experience to identify certain genuine qualities and connections, namely, phenomenal qualities and connections. (E.g., we can experience green but not grue. If we were unable to identify phenomenal qualities and connections in this way, we could not notice change or constancy in our experience, nor could we even identify so-called recalcitrant experiences.) Having done this, we may then seek causal explanations of why we experience the particular phenomenal qualities and connections that we do. Among the competing explanations, consider those that posit theoretical qualities and connections described solely in terms of known phenomenal qualities and connections, the concept of causation, the general concepts of quality and connection, and any other transcendentally justified concepts (i.e., any other concepts that are required in order to engage in theory construction at all). Since these explanations are all formulated with the same terms, one can straightforwardly compare their complexity without running into the relativist's worry that Cambridge properties and relations might sneak in under the veil of a superficially simple syntax of primitive theoretical terms. After doing this, one would be justified in identifying the simplest of these explanations as correct. Then, from this explanation one can extract an authoritative list of theoretical qualities and connections. Such a list would bring one a long way toward a solution to the epistemological problem. Suppose, however, that this procedure should fail to isolate a unique causal explanation—and, hence, a unique list of theoretical qualities and connections. The resulting situation would not be revolutionary; it would be just one more instance of the familiar problem of the underdetermination of theory by the data.

Given that qualities and connections form a special category of properties and relations, which of the two traditional conceptions of properties and relations applies to this special category? Consider an example involving shape. Take the following little object:

(1)



What shape is figure (1)? In answer to this question, one might say that (1) is triangular. Or one could equally well say that (1) is trilateral. Each of these answers suffices to inform us of its shape. The reason for this is that, intuitively, the quality of being triangular and the quality of being trilateral are the very same quality. They are how it is with (1) in regard to shape. Though the concept of being triangular and the concept of being trilateral are distinct, they correspond to the same quality of things in the world. Indeed, there is no limit to the number of necessarily equivalent ways to conceive of this shape. Yet there is only one shape. On conception 1, necessary equivalence is sufficient for identity while on conception 2 it is not. It would seem, therefore, that qualities and connections, including this shape, conform to conception 1 whereas concepts conform instead to conception 2. Qualities and connections are what fix the actual conditions in the world, and, as such, they do not exhibit distinctions finer than necessary equivalents. Concepts, on the other hand, pertain primarily to thinking about the world; it is in thinking that finer intensional distinctions show up.

Now consider the sort of things called *conditions*. Conditions are the sort of things that are said to obtain (or to be so). For example, the condition that (1) is triangular obtains, and the condition that (1) is circular does not obtain. Similarly, the condition that something is triangular obtains, and the condition that nothing is triangular does not. The conditions that obtain constitute, as we say, how it is in the world. They are the actual states of affairs.

Just as qualities and connections conform to conception 1, so do conditions. (Of course, conditions are 0-ary whereas qualities are 1-ary and connections are n-ary, $n \geq 2$.) To see this, consider again the example of figure (1). Intuitively, the condition that (1) is triangular is the same condition as the condition that (1) is trilateral. Indeed, the condition that (1) is triangular is intuitively the same condition as the condition that (1) is trilateral and $5 + 7 = 12$. And so on for all necessarily equivalent conditions. Conditions, like other non-intentional things, do not exhibit distinctions finer than necessary equivalents.

But how are qualities, connections, and conditions connected to one another? What for example is the connection between the quality of curving and the little figure (2) and the condition that (2) curves?

(2)

The answer is that the condition that (2) curves is the result of predicating the quality of curving of figure (2). And what is the connection between the quality of curving and the condition that something curves? The answer is that the condition that something curves is the existential generalization of the quality of curving. Likewise, the condition that (1) is triangular and (2) curves is the conjunction of the condition that (1) is triangular and the condition that (2) curves; and so on. Let us call the fundamental logical operations like these *condition-building operations*. With this terminology one may then say generally how conditions are related to qualities and connections. Conditions are built up by means of condition-building operations from qualities and connections,[10] i.e., from the properties and relations that provide the world with its logical, causal, and phenomenal order. Of course, since conditions conform to conception 1, there are any number of ways in which a given condition can be built up. Witness the identity of, e.g., the condition that (1) has three angles, the condition that (1) has three sides, the condition that (1) has three sides and $5 + 7 = 12$, and so on.

This completes my informal characterization of qualities, connections, and conditions. Let us now turn our attention to intentional matters and thinking.

41. Thoughts and Concepts

Ideas, Locke tells us, are '...*whatever it is which the mind can be employed about in thinking*...' (p. 32, *An Essay Concerning Human Understanding*). He continues, '...there are such *ideas* in men's minds: every one is conscious of them in himself; and men's words and actions will satisfy him that they are in others.' Ideas divide naturally into two kinds, *thoughts* and *concepts*. Let us consider thoughts first. Just as conditions are the sort of things that are said to obtain or not, so thoughts are the sort of things that are said to be true or false. According to common sense, a thought is true if and only if it corresponds to a condition that obtains, and a thought is false if and only if it corresponds to a condition that does not obtain. For example, the thought that something is triangular is true just in case the corresponding condition that something is triangular obtains. And the thought that nothing is triangular is false just in case the corresponding condition that nothing is triangular does not obtain.

I come next to concepts. Just as qualities are said to qualify objects and connections are said to connect objects, concepts are said to apply to objects. For example, just as the quality triangular qualifies figure (1), the concept of being triangular applies to figure (1). There are, relatedly, concepts that correspond to qualities or connections. For example, the concept of being triangular corresponds to the quality of being triangular. If a concept corresponds to a quality, then the concept applies to an object if and only if the corresponding quality qualifies the object. Or if a concept corresponds to a connection, then the concept applies to certain objects (in a certain order) if and only if the corresponding connection connects those objects (in that order).[11]

Thoughts are the sort of thing that can be believed, disbelieved, remembered, forgotten, understood, misunderstood, asserted or denied in language, advanced as theories, etc. This is to say, thoughts are natural objects of intentional relations. Now given my investigation into the logic for intentional matters, it follows that these objects of intentional relations conform to conception 2. This makes thoughts quite a different type of thing from conditions. Consider a few examples involving the little triangular figure (1) discussed earlier. Even though the condition that (1) is triangular is the same condition as the condition that (1) is trilateral, the thought that (1) is triangular is, intuitively, quite distinct from the thought that (1) is trilateral. And of course, even though the condition that (1) is triangular is the same condition as the condition that (1) is triangular and $5 + 7 = 12$, the thought that (1) is triangular is quite distinct from the thought that (1) is triangular and $5 + 7 = 12$. And so on. Similar considerations show that concepts too are conception 2 intensions. Therefore, thoughts, since they can be said to be true or false, are 0-ary conception 2 intensions, and concepts, since they can be said to apply to objects, are n-ary conception 2 intensions, $n \geq 1$.[12]

But there is a lacuna in this story: where do the Cambridge properties and relations fit in? Consider the Cambridge property grue. Since grue is not a phenomenal property, we cannot experience it. And since it has no causal efficacy, it can leave no causal traces in the world. Nor is it a fundamental logical property. Indeed, our only knowledge of grue comes via the original definition: x is grue *iff*$_{\mathrm{df}}$ x is green if examined before t and blue otherwise. However, by the conclusion arrived at near the close of §38, complex expressions in natural language express the kind of intensions that have to do with intentionality and thinking. That is, they express thoughts and concepts. Therefore, the complex expression 'green if examined before *t* and blue otherwise' expresses a concept. What could grue be if it is not this concept? True, there is one alternative; someone might think that grue is one of the conception 1 properties posited in the theory T1. To be sure, T1 does posit such a property since T1-property formation is closed under all combinations of the logical operations, even *ad hoc* ones. But the fact that T1 posits such properties does not decide the question of whether they really exist. For T1 is only a provisional theory which was constructed at a precritical, experimental stage, and the relevant closure property was built in to simplify the construction rather than to capture a philosophically motivated picture of intensional entities. Thus, T1 cannot be used to settle the present basic philosophical issue. And once one sets aside T1 as an authority, one sees that the most natural and economical picture is that in which grue is simply identified with the concept expressed in the original definition. What good reason could there conceivably be for identifying grue with anything but this concept? I conclude, therefore, that grue is a mere concept. And generalizing on this, I conclude that all Cambridge properties and relations are nothing but concepts. That is, all properties that are not qualities and all relations that are not connections are concepts. After all, these properties and relations play no role in determining the logical, causal, or phenomenal order in the world. Their primary role is in thinking, and that role is often only playful. Indeed, grue was a concept introduced with no other purpose than the posing of a riddle.

I thus arrive at a natural and economical picture. Qualities, connections, and conditions are the intensional entities that pertain to the world. Thoughts and concepts are those that pertain to thinking. And qualities, connections, conditions, thoughts, and concepts are all the intensional entities there are.

The next question to consider is how these intensional entities are related to one another. Recall the little curving figure (2) discussed earlier. What is the connection between the quality of curving, figure (2), and the thought that (2) curves? Just as the condition that (2) curves is a result of predicating curving of (2), so too the thought that (2) curves is a result of predicating curving of (2). But the thought that (2) curves is quite distinct from the corresponding condition that (2) curves. It follows that two different types of predication must be at work here: one combines the quality of curving and figure (2) to yield the thought that (2) curves, and the other combines the quality of curving and figure (2) to yield the condition that (2) curves. Likewise for other cases. For example, just as the condition that something curves is a result of existentially generalizing on the quality of curving, so too the thought that something curves is a result of existentially generalizing on this quality. Since this thought is quite distinct from the corresponding condition, two types of existential generalization must be at work. One operates on the quality of curving to yield the condition that something curves, and the other operates on this quality to yield the thought that something curves. In this way I isolate two distinct types of fundamental logical operations—*condition-building operations* and *thought-building operations*. This leads to the following picture. Thoughts and conditions are alike except that, whereas conditions are built up ultimately from qualities and connections by means of condition-building operations, thoughts are built up ultimately from qualities and connections by means of thought-building operations. Of course, since conditions conform to conception 1, they can be built up in any number of ways; since thoughts conform to conception 2, they are built up in a unique, non-circular way.

Notice, however, that concepts can themselves be combined together to obtain thoughts. This suggests extending the above picture to allow that concepts also are built up ultimately from qualities and connections by means of the thought-building operations. Or to put the point the other way around, concepts can be analysed ultimately into qualities and connections by means of the inverses of the thought-building operations. A consequence of this picture is that there are certain concepts that, as limiting cases, cannot be analysed any further by means of the inverses of the thought-building operations. These concepts I will call *simple concepts* or, following Locke, *simple ideas*. All other ideas, whether they be thoughts or concepts, I will call *complex*. Since simple concepts cannot be analysed any further, according to the above picture they must be just qualities or connections. For example, the concept of green is just the quality green; the concept of predication is just the connection predication; etc. In turn, since simple concepts are qualities and connections, they must conform to conception 1. That is, simple concepts are identical if and only if they are necessarily equivalent. Complex ideas are not like this; they can differ even if they are necessarily equivalent.

This metaphysical picture of the constitution of thoughts and concepts is really just the outcome of conclusions reached earlier in the book. At the close of §38 I argued that sentences in natural language express thoughts. However, we know that many sentences in natural language are used to express theories and phenomenal descriptions. Therefore, theories and phenomenal descriptions are simply kinds of thoughts. Given the special roles that qualities and connections play in theories and phenomenal descriptions, it follows that genuine qualities and connections (plus perhaps subjects of singular predications) must be the ultimate building blocks of all true theoretical and phenomenal thoughts. Moreover, an analogous argument shows that genuine qualities and connections (plus perhaps subjects of singular predications) must be the ultimate building blocks of all the concepts that go to make up true theoretical and phenomenal thoughts. The metaphysical picture sketched above is nothing but the generalization on these conclusions: qualities and connections (plus subjects of singular predications) are the ultimate building blocks of all thoughts and concepts. In view of the unique foundational role that theoretical and phenomenal concepts have in all other thoughts and concepts, this generalization is compelling.[13]

I have been unable to find any counterexamples to the above theory, and I have some grounds for believing that none will be forthcoming. Any purported counterexample that someone produces must be a thought or concept to which he has some kind of epistemic access. Specifically, he must know of the thought or concept either innately or by experience or by description—such as causal "reference-fixing" description—or by definition, using the fundamental condition-building and thought-building operations. However, the metaphysical theory sketched in this chapter is designed to account for all intensional entities to which we have these kinds of epistemic access, and thus, counterexamples should not be forthcoming.

42. Realism and Representationalism

The foregoing theory that the primary bearers of truth (i.e., thoughts) are built up ultimately from the primary constituents of reality is by no means novel. It harks back to views held by Plato and Aristotle. According to Plato, those things that can be said to be true or false 'owe their existence to the weaving together of forms' (*Sophist* 260E). Likewise, according to Aristotle, 'Nothing, in fact, that is said without combination [literally, without interweaving] is either true or false' (*Categories* 13^b11), and the primary constituents of reality, i.e., the items in primary metaphysical categories, are those 'things said without any combination', i.e., literally, those things said without any interweaving (*Categories* 1^b25). This similarity is not superficial. Recall that Plato and Aristotle are usually identified as the originators of the correspondence theory of truth. But what is the relation of *correspondence*? Given the above theory of the constitution of thoughts and concepts, the correspondence relation can be given a precise logical analysis, an analysis that is plainly implicit in this metaphor of interweaving invoked by Plato and Aristotle.

I will take a moment now to show how this analysis of the correspondence relation will go. The goal is to say what it is for a thought to correspond to a condition and what it is for a concept to correspond to a quality or a connection. Consider once again the little curving figure (2) discussed earlier. Why does the thought that (2) curves correspond to the condition that (2) curves? The answer is that the thought that (2) curves and the condition that (2) curves are formed ("woven together") in the same way from the same basic things, the only difference being that the thought is formed by means of the thought-building operation of predication whereas the condition is formed by means of the condition-building operation of predication. The thought that (2) curves is true because the condition to which it bears this structural isomorphism is a condition that obtains. For another example consider the shape of the little figure (1) which I discussed earlier. To this quality there correspond any number of necessarily equivalent concepts, e.g., the concept of being triangular, the concept of being trilateral, the concept of being a closed figure whose angles sum to two right angles, the concept of being triangular and such that $5 + 7 = 12$, etc. These concepts all correspond to the single quality triangular. But why? When these concepts are analysed by means of the inverses of the thought-building operations and then formed again, this time by means of the condition-building operations, the result is just the shape of (1), i.e., the quality triangular itself. Thus, in general, the relation of correspondence that holds between an idea and a quality, connection, or condition is the relation holding between entities that are composed in the same way from the same ultimate constituents, the only difference being that the one is composed by means of the fundamental logical operations typically used for composing thoughts whereas the other is composed by means of the fundamental logical operations typically used for composing conditions. It is in this structural isomorphism that we find a purely logical analysis of the relation of correspondence.

Before showing how to make this analysis fully precise and rigorous, I want to take up the general philosophical question of how an idea is fundamentally related to a counterpart in the world. One can discern two opposing trends in the history of philosophy concerning this question, one that may be called *realism* and the other, *representationalism*. Take any complex idea that corresponds to a quality, connection, or condition. Since the idea is never identical to the thing to which it corresponds, there is a sense in which it can be said to represent. This is not what I mean by representationalism. For me, representationalism is the much stronger doctrine that, even when an idea is fully analysed, neither qualities nor connections nor items belonging to other primary metaphysical categories (such as particulars, stuffs, etc.) ever enter in. A consequence of representationalism is that there is no way to escape representation; at most, thoughts and concepts give way to other thoughts and concepts, *ad infinitum*. The effect of this is to render the link between ideas and the world some kind of mysterious unanalysable relation. Realism, on the other hand, is the doctrine that, if an idea is fully analysed, then qualities, connections, and perhaps items from other primary metaphysical categories do enter in. So according to realism, representation always comes to a halt at some stage; sooner or later one gets to the real things in the world, to the primary constituents of reality. In this way realism opens up the possibility of giving a non-circular analysis of how ideas are linked to the things in the world to which they correspond. The theory that I have described, for example, is realistic since all ideas can be analysed ultimately into qualities and connections and perhaps subjects of singular predication.

The history of philosophy is laced with conflicting representationalist and realist threads. In fact, there is often evidence of both doctrines in the work of a single philosopher. Nevertheless, as I have already intimated, one may venture to classify Plato and Aristotle as realists. And in contrast, one may venture to classify as representationalists nearly all of the classical modern philosophers from Descartes and Locke through Kant. Further, in classical modern logical philosophy one may classify Russell as a realist and Frege as a representationalist.[14] In contemporary philosophy in the English-speaking world there are currents of neo-Russellian realism; nonetheless, representationalism still exercises remarkable influence.[15] And in continental European philosophy the tradition of representationalism has been all but continuous from Descartes to Derrida.

Yet despite its long history, representationalism is an inherently defective doctrine. The reason for this could scarcely be more dramatic: representationalism inevitably veils its own subject matter—thought and language—in mystery and metaphor. So if realism can as much as be made acceptable, it must be counted as superior to representationalism. I submit, however, that the realism that I have set forth is perfectly acceptable. The only remaining task is to show that it can be given a coherent formal statement, and this will be done in the next section. My conclusion is that there is just no good reason to remain under the spell of representationalism.

In closing I should like to add that the virtue of realism is not limited to its ability to clear up the intellectual mist produced by representationalism. Realism also leads to promising solutions to some of the most central outstanding problems in classical modern philosophy, problems that have resisted solution for so long largely because they have been thought of in representationalist terms. These problems and their realistic solutions will be the final concern of the work.

43. The Logic for Qualities and Concepts*

The version of realism I am advocating is a synthesis of the two traditional conceptions of intensional entities. To adopt it, I must make certain revisions in the two provisional theories that I have been working with in the preceding chapters. These revisions are concerned for the most part with the theory for conception 1.

* Readers seeking only an overview may proceed to the next chapter.

On the suggested version of realism, there are two types of intensional entities: qualities, connections, and conditions, on the one hand, and thoughts and complex concepts, on the other. Qualities, connections, and conditions conform to conception 1 while thoughts and complex concepts conform to conception 2. Since these are the only intensional entities, properties and relations fall into two kinds: those that are genuine qualities and connections and those that are mere complex concepts. The latter are the Cambridge properties and relations. The Cambridge property grue, for example, is a complex concept and so is a conception 2 entity. On the provisional theory for conception 1, however, no matter how conception 1 entities are combined together by means of the fundamental condition-building operations, the result is always treated as another conception 1 entity. Thus, on that provisional theory, grue would be counted as a conception 1 entity; that is, it would be counted as a quality, not as a complex concept. To eliminate this conflict, I must modify the characterization of the condition-building operations presented in the provisional theory.

Consider some simple examples to see how this modified characterization should go. Take the condition-building operation of negation. As in the provisional theory, this operation should still take, e.g., the condition that something is green to the condition that nothing is green. But now this operation should take the quality green to the concept not-green. The reason for this is that there is no quality not-green, and therefore, the property not-green must be a concept, namely, the concept not-green. At the same time, the condition-building operation of negation should take this property not-green back to the quality green. The reason is that the property not-not-green is necessarily equivalent to the quality green, and therefore, this property must be a quality, namely, the quality green. In general, the condition-building operations must accord with the following principle: if a property, relation, or condition is necessarily equivalent to a quality, connection, or condition, then it is identical to that quality, connection, or condition; otherwise it is just an appropriate concept.

In §41 I divided ideas into two kinds—thoughts and concepts. A complex idea was defined as an idea that can be analysed by means of the inverses of the thought-building operations; a simple idea, as one that cannot. This made every simple idea either a quality or a connection. In what follows, however, I will simplify things by also permitting conditions to be simple ideas. In consequence, every intensional entity will be called an idea.

With the foregoing in mind, I will now construct a new type of model structure, one designed to model the behavior of qualities, connections, conditions, thoughts, and concepts. Thus, I define a *type 3 model structure* $\mathscr{M}$ to be any structure

$$\begin{aligned}&\langle \mathscr{D}, \mathscr{P}, \mathscr{K}, \mathscr{G}, \mathrm{Id},\\ &\mathrm{Conj}^c, \mathrm{Neg}^c, \mathrm{Exist}^c, \mathrm{Exp}^c, \mathrm{Inv}^c, \mathrm{Conv}^c, \mathrm{Ref}^c, \mathrm{Pred}_0^c, \mathrm{Pred}_1^c, \ldots,\\ &\mathrm{Conj}^t, \mathrm{Neg}^t, \mathrm{Exist}^t, \mathrm{Exp}^t, \mathrm{Inv}^t, \mathrm{Conv}^t, \mathrm{Ref}^t, \mathrm{Pred}_0^t, \mathrm{Pred}_1^t, \ldots\rangle\end{aligned}$$

that satisfies the following three requirements. Let the diminished structure

$$\begin{aligned}&\langle \mathscr{D}, \mathscr{P}, \mathscr{K}, \mathscr{G}, \mathrm{Id}, \mathrm{Conj}^c, \mathrm{Neg}^c, \mathrm{Exist}^c, \mathrm{Exp}^c,\\ &\mathrm{Inv}^c, \mathrm{Conv}^c, \mathrm{Ref}^c, \mathrm{Pred}_0^c, \mathrm{Pred}_1^c, \ldots\rangle\end{aligned}$$

be called $\mathscr{M}_1$, and let the diminished structure

$$\begin{aligned}&\langle \mathscr{D}, \mathscr{P}, \mathscr{K}, \mathscr{G}, \mathrm{Id}, \mathrm{Conj}^t, \mathrm{Neg}^t, \mathrm{Exist}^t, \mathrm{Exp}^t,\\ &\mathrm{Inv}^t, \mathrm{Conv}^t, \mathrm{Ref}^t, \mathrm{Pred}_0^t, \mathrm{Pred}_1^t, \ldots\rangle\end{aligned}$$

be called $\mathscr{M}_2$. The first requirement on $\mathscr{M}$ is simply that the diminished structures $\mathscr{M}_1$ and $\mathscr{M}_2$ are standard algebraic model structures. Before I give the second requirement I will give some definitions. Let the operations $\mathrm{Conj}^c, \ldots$ be called condition-building operations, and let $\mathrm{Conj}^t, \ldots$ be called thought-building operations. Things in a subdomain $\mathscr{D}_i$, for $i \geq 0$, are called ideas. Things in a $\mathscr{D}_i$, for $i \geq 1$, are called concepts. Ideas that are in the range of some thought-building operation are called complex. And ideas that are not complex are called simple. Complex ideas in $\mathscr{D}_0$ are called thoughts. Simple ideas in $\mathscr{D}_1$ are called qualities. And simple ideas in a $\mathscr{D}_i$, for $i \geq 2$, are called connections. Things in $\mathscr{D}_0$ that are in the range of some condition-building operation are called conditions. Things in $\mathscr{D}_1$ that are in the range of a condition-building operation are called properties. Things in a $\mathscr{D}_i$, for $i \geq 2$, that are in the range of some condition-building operation are called relations. Properties and relations that are complex ideas are called Cambridge properties and relations. Now for the second requirement on $\mathscr{M}$. This requirement concerns the type that the algebraic model structures $\mathscr{M}_1$ and $\mathscr{M}_2$ must be. $\mathscr{M}_2$ is type 2. $\mathscr{M}_1$ however is mixed. In the case of qualities, connections, and conditions, $\mathscr{M}_1$ behaves like a type 1 model structure, but when it comes to Cambridge properties and relations, $\mathscr{M}_1$ behaves like a type 2 model structure. Specifically, if x is a quality, connection, or condition, then for any property, relation, or condition y, x and y satisfy the following: $(\forall H \in \mathscr{H})(H(x) = H(y)) \supset x = y$. On the other hand, if a property or relation $\mathrm{Conj}^c(u, v)$ is a Cambridge property or relation, then it is identical to the complex concept $\mathrm{Conj}^t(u, v)$; if a property or relation $\mathrm{Neg}^c(u)$ is a Cambridge property or relation, then it is identical to the complex concept $\mathrm{Neg}^t(u)$; and so on *mutatis mutandis* for each of the other condition-building operations. Before I give the third requirement on $\mathscr{M}$, I will give one more definition. Take any element of $\mathscr{D}$ that is built up from elements of $\mathscr{D}$ by means of the condition-building and thought-building operations. Consider the tree associated with this building-up procedure. If in this tree a given node branches into the three nodes $\langle \mathrm{Pred}^c_k, v, w\rangle$ or $\langle \mathrm{Pred}^t_k, v, w\rangle$, then the node w will be called a *subject node*. The third requirement on $\mathscr{M}$ concerns the constitution of simple and complex ideas. First, every condition is a simple idea. Secondly, every simple idea has an associated tree (infinitely many, in fact) in which every terminal node is either a condition-building operation, a quality, a connection, or a subject node. (Such a tree will be called a *condition-building tree*.) In turn, every complex idea has an associated tree in which every terminal node is either a thought-building operation, a simple idea, or a subject node.[16] (Such a tree will be called a *thought-building tree*. For convenience I will also say that a simple idea is as a limiting case a one-node condition-building tree for itself and a one-node thought-building tree for itself.)

With type 3 model structures defined I can now go on to develop the logic for qualities and concepts. The strategy will be to proceed in stages akin to those encountered over the previous chapters. Thus, I begin by constructing an intensional language $\mathscr{L}_\omega$ appropriate to the logic for qualities and concepts. In its primitive symbols $\mathscr{L}_\omega$ is just like L_ω except that it contains an additional punctuation mark $|$. The terms and formulas of $\mathscr{L}_\omega$ are inductively defined as follows:

(1) All variables are terms.
(2) If $t_1, \ldots, t_j$ are terms, $F^j_i(t_1, \ldots, t_j)$ is a formula.
(3) If A and B are formulas and v_k is a variable, then $(A \,\&\, B)$, $\neg A$, and $(\exists v_k)A$ are formulas.

(4) If A is a formula and $v_1, \ldots, v_m$ are distinct variables, for $m \geq 0$, then $|A|_{v_1 \ldots v_m}$ and $[A]_{v_1 \ldots v_m}$ are terms.

The complex singular terms in $\mathscr{L}_\omega$ should be read as follows ($m \geq 2$ and $n \geq 1$): $|A|$, the condition that A; $|A|_{v_1}$, the property of things v_1 such that A; $|A|_{v_1 \ldots v_m}$, the relation among things $v_1, \ldots, v_m$ such that A; $[A]$, the thought that A; $[A]_{v_1 \ldots v_n}$, the concept of things $v_1, \ldots, v_n$ such that A.[17]

In constructing the semantics for $\mathscr{L}_\omega$ I will make use of the following heuristic principles. Qualities and connections are the only conception 1 properties and relations. All other properties and relations, since they are Cambridge properties and relations, are complex concepts. So, in particular, if there exists a quality of (or connection among) things α such that A, then that quality (connection) is identical to the property of (relation among) things α such that A. On the other hand, if there does not exist a quality of (connection among) things α such that A, then the property of (relation among) things α such that A is a Cambridge property (relation) and, hence, is just identical to the concept of things α such that A. The relevant semantics for $\mathscr{L}_\omega$ may be constructed as follows.

The notions of interpretation, assignment, truth, and validity for $\mathscr{L}_\omega$ (relative of course to type 3 model structures) are defined exactly as they are in the §14 semantics for L_ω. When I come to the definition of the denotation function, however, a few alterations are in order. The denotation function $\mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}$ for $\mathscr{L}_\omega$ must be defined for all terms, including the two types of intensional abstracts $|A|_\alpha$ and $[A]_\alpha$. Accordingly, although the clauses for variables and elementary intensional abstracts $[F_i^j(v_1, \ldots, v_j)]_{v_1 \ldots v_j}$ are unchanged, the clauses for the non-elementary intensional abstracts $[A]_\alpha$ are modified in two ways. First, it must be made explicit that the fundamental logical operations are the thought-building operations $\mathrm{Conj}^{\mathrm{t}}, \ldots$. Secondly, the clause for predication$_k$, $k \geq 1$, now covers, not only k-ary relativized predications of the form $[F_i^j(t_1, \ldots, t_{h-1}, [B]_\gamma^\delta, t_{h+1}, \ldots, t_j)]_\alpha$, but also ones of the form $[F_i^j(t_1, \ldots, t_{h-1}, |B|_\gamma^\delta, t_{h+1}, \ldots, t_j)]_\alpha$. Finally, the following clause is added for the complex terms $|A|_\alpha$:

If relative to $\mathscr{M}$ there is a quality, connection, or condition x in the same subdomain as $\mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}([A]_\alpha)$ and if $(\forall H \in \mathscr{K})(H(x) =$

$H(\mathbf{D}_{\mathcal{I}\mathcal{A}\mathcal{M}}([A]_\alpha)))$, then $\mathbf{D}_{\mathcal{I}\mathcal{A}\mathcal{M}}(|A|_\alpha) = x$; otherwise $\mathbf{D}_{\mathcal{I}\mathcal{A}\mathcal{M}}(|A|_\alpha) = \mathbf{D}_{\mathcal{I}\mathcal{A}\mathcal{M}}([A]_\alpha)$.

Relative to this semantics, a logic T3 for $\mathscr{L}_\omega$ may be formulated, and I am optimistic that there is a positive solution to the completeness problem for it. In this connection, $\Box$ and $\Diamond$ may, as a notational convenience, be defined as follows:

$\Box A$ *iff*$_{\mathrm{df}}$ $|A| = ||A| = |A||$
$\Diamond A$ *iff*$_{\mathrm{df}}$ $\neg\Box\neg A$.

Thus, just as in T1, so also in T3, modal logic may be viewed as the identity theory for intensional abstracts.

Now recall the extensional analysis of intensional abstraction given in §37. There it was shown how to translate the intensional language L_ω into an extensional language L. By a fully analogous procedure, the intensional language $\mathscr{L}_\omega$ for the logic of qualities and concepts can be translated into an extensional language $\mathscr{L}$. Among the primitive logical constants of $\mathscr{L}$ are distinguished logical predicates in terms of which one can express the condition-building and thought-building operations. In $\mathscr{L}$, therefore, one can define the concepts of quality, connection, condition, simple and complex idea, thought, concept, Cambridge property, and Cambridge relation.

Finally, at any stage in the study of the logic for qualities and connections, Δ may be singled out as a distinguished logical predicate for the predication relation, and appropriate axioms for it may be introduced.

44. Correspondence

Earlier I gave an informal statement of a purely logical definition of the correspondence relation. According to the definition an idea x corresponds to a quality, connection, or condition y if and only if x and y can be composed from the same entities in the same way, the only difference being that x is composed by means of the thought-building operations whereas y is composed by means of the condition-building operations. Thus, correspondence turns out to be a structural relation, specifically, a structural isomorphism. In this, correspondence is not a mysterious, unanalysable relation of ideas "mirroring" the world. For that matter, the definition allows us to say precisely what it is about certain ideas that prompts us to say that they "mirror" the world. I will have more to say about this in a moment. First, let me show what the analysis looks like formally.

I will begin by giving a model-theoretic presentation. Consider any type 3 model structure $\mathscr{M}$. Relative to $\mathscr{M}$, a thought-building tree t is defined to be *isomorphic* to a condition-building tree t' if and only if t and t' are exactly alike in all their terminal nodes except that, when a condition-building operation occurs as an operation (not as an argument) at some terminal node in t', the equivalent thought-building operation occurs in its place at that node in t.[18] Then the correspondence relation on $\mathscr{D}$ relative to $\mathscr{M}$ is defined as follows: an idea $x \in \mathscr{D}$ *corresponds* to a quality, connection, or condition $y \in \mathscr{D}$ if and only if x has a thought-building tree that is isomorphic to a condition-building tree of y.[19] (Of course, since y is a conception 1 entity, it has infinitely many condition-building trees.)

This definition of the correspondence relation for $\mathscr{M}$ is justified by the following theorem:

> *Theorem:* Let $\mathscr{M}$ be any type 3 model structure. Then for all ideas x and all qualities, connections, and conditions y in the domain of $\mathscr{M}$, x corresponds to y if and only if x and y belong to the same subdomain and x and y are necessarily equivalent in $\mathscr{M}$ (i.e., $x, y \in \mathscr{D}_i$, for some $i \geq 0$, and $(\forall H \in \mathscr{K})H(x) = H(y)$).

(The proof is a straightforward inductive argument.) It is now clear why the correspondence relation is entitled to play such an important role in linking ideas to the world. Correspondence is a relation holding between ideas and necessarily equivalent qualities, connections, and conditions.

But I have still not defined the correspondence relation itself. I have only defined a relation on elements of the domain of arbitrary type 3 model structures $\mathscr{M}$. Yet since this is only a construction in model theory, these elements could be any arbitrary objects whatsoever; they need not even be intensional entities. The next step, therefore, is to define the correspondence relation proper. To do this, I will use the canonical language $\mathscr{L}$ for the logic of qualities and concepts, where this language is interpreted in the intended way. After all, at some stage it is appropriate to kick away the metatheoretical study aids and to take one's logical theory at face value. That is, at some stage it is appropriate to stop mentioning the language of one's logical theory and instead to begin using it. Now is the time.

Recall that in $\mathscr{L}$ one can express the thought-building and condition-building operations. Thus, on analogy with the model-theoretic definition one can in $\mathscr{L}$ give an inductive definition of the correspondence relation. I.e., one can give an inductive definition of the relation holding between ideas x and qualities, connections, and conditions y such that x has a thought-building tree that is isomorphic to one of the condition-building trees of y. The next step is to single out in $\mathscr{L}$ the distinguished logical predicate Δ for the predication relation. Then, given a Zermelo-style theory for Δ, one can turn the inductive definition of the correspondence relation into a direct definition. This is so since the relevant thought-building and condition-building trees have only a finite number of nodes. Thus, using the canonical language $\mathscr{L}$, one obtains a purely logical direct definition of the correspondence relation.

According to the definition, correspondence is a structural relation. Specifically, it is structural isomorphism definable within pure logic in terms of the predication relation and the fundamental operations by means of which thoughts and conditions are composed. Given the logical behavior of these fundamental logical relations, the structural isomorphism insures that if an idea corresponds to a quality, connection, or condition, then they are necessarily equivalent.

Correspondence is thus not a mysterious, unanalysable extra-logical relation of ideas "mirroring" the world. In fact, one can finally say precisely what it is about certain ideas that makes us want to say that they "mirror" the world. An idea "mirrors" a quality, connection, or condition in the world if and only if the two can be composed in exactly the same way from exactly the same things except that in the case of the idea the thought-building operations assume the role played by the condition-building operations. In this fashion the analysis of the correspondence relation preserves—and indeed makes good logical sense of—the metaphor that ideas "mirror" the world. In fact, the analysis does this with a flourish. Recall that qualities, connections, and conditions conform to conception 1, and ideas conform to conception 2. This shows that there exist infinitely many distinct yet perfectly faithful "mirrors" of any single quality, connection, or condition in the world.

9

Logic

45. Truth

The ancient problem of truth was to say informally what truth is. That problem was solved by Plato and Aristotle by means of their germinal versions of the correspondence theory. The modern problem of truth is not to seek a novel definition; rather, it is to find a logically precise and clear expression of the ancient Greek definition. This at any rate is how Alfred Tarski sees the modern problem:

> We should like our definition to do justice to the intuitions which adhere to the *classical Aristotelian conception of truth*—intuitions which find their expression in the well-known words of Aristotle's *Metaphysics*:
>
> > 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.
>
> However, all these formulations [i.e., the original Aristotelian formulation and the sundry modern attempts to capture it] can lead to various misunderstandings, for none of them is sufficiently precise and clear . . .; at any rate none of them can be considered a satisfactory definition of truth. It is up to us to look for a more precise expression of our intuitions. (§3, 'The Semantic Conception of Truth and the Foundations of Semantics')

Tarski's semantic conception of truth is the backbone of the Tarskian technique for characterizing validity for formal languages. I will show in a while (§47) that there are significant obstacles to a general account of validity along Tarski's semantic lines. But first let me examine the semantic conception of truth, not as a formal device used in the model-theoretic characterization of validity, but rather as a serious account of truth itself. This exercise will help to motivate my own solution to the modern problem of truth.

I gave evidence in §6 and §8 for the thesis that 'that'-clauses are singular terms whose semantical correlates are propositions, rather than linguistic entities such as sentences. Given this thesis, there is good evidence that truth is a property of propositions, as is suggested by the following intuitive validity containing a 'that'-clause:

It is true *that snow is white* if and only if snow is white.

Given the theory of qualities and concepts advanced in the previous chapter, the type of propositions that can have the property of truth are the ones known as *thoughts*. Therefore, I conclude that, as a minimum requirement, an adequate theory of truth must provide an account of what it takes for a thought to be true.

Now despite the fact that truth is a property of thoughts, we also commonly say of *sentences* that they are true relative to some language or other. For example, we say that English sentences are true (i.e., true-in-English), that French sentences are true (i.e., true-in-French), that Polish sentences are true (i.e., true-in-Polish), and so on for all languages. The semantic conception is predicated on this elementary linguistic fact.

The word 'true' thus has a multiplicity of uses—one use for propositions and an indefinitely large constellation of uses for sentences, one distinct use for each of the various languages. How are we to explain this multiplicity of uses?

One explanation is that it is merely ambiguity, merely *chance homonymy*. Tarski's polemical remarks about the 'meaninglessness' of 'those endless, often violent discussions on the subject: "What is the right conception of truth?"' and his advice that we should 'reconcile ourselves with the fact that we are confronted, not with one concept, but with several different concepts which are denoted by the same word' (§14, 'The Semantic Conception') suggest that he is sympathetic with this chance homonymy explanation. This suggestion is borne out by Tarski's strategy for giving semantic truth definitions. For he holds that,

> ...we must always relate the notion of truth, like that of a sentence, to a specific language; for it is obvious that the same expression which is a true sentence in one language can be false or meaningless in another. (§2, 'The Semantic Conception')

This leads him to define the concepts of true-in-*L*, for different languages *L*, wholly independently of one another.

There is, however, another explanation of the multiplicity of uses of 'true', one which is much better. It is an application of Aristotle's theory of non-chance homonymy, or *focal meaning*, as it has been called.[1] Aristotle gives focal-meaning accounts for multiple uses of 'be', 'healthy', and 'medical':

> There are many senses in which a thing may be said to 'be', but all that 'is' is related to one central point, one definite kind of thing, and is not said to 'be' by a mere ambiguity. Everything which is healthy is related to health, one thing in the sense that it preserves health, another in the sense that it produces it, another in the sense that it is a symptom of health, another because it is capable of it. And that which is medical is relative to the medical art, one thing being called medical because it possesses it, another because it is naturally adapted to it, another because it is a function of the medical art. And we shall find other words used similarly to these.
>
> As, then, there is one science which deals with all healthy things, the same applies in the other cases also. For not only in the case of things which have one common notion does the investigation belong to one science, but also in the case of things which are related to one common nature; for even these in a sense have one common notion.
>
> But everywhere science deals chiefly with that which is primary, and on which the other things depend, and in virtue of which they get their names. (*Metaphysics*, book Gamma, 1003^{a}3–1003^{b}18)

A focal-meaning explanation of the multiple uses of 'true' goes as follows. The fact that English sentences can be said to be true in their way (i.e., true-in-English), that French sentences can be said to be true in another (i.e., true-in-French), etc. is not a matter of a "mere ambiguity". Rather they are said to be true in virtue of the fact that they are all related to "one central point", namely, a central concept of truth. Specifically, for each given language L, a sentence is true-in-L if and only if what it expresses in L is true in the primary sense. Since what a sentence expresses is a thought, thoughts are the things that are true primarily. So the central concept of truth is the concept of a true thought. Sentences in the various languages are called true only secondarily, by virtue of the fact that they are related via their respective meaning relations to true thoughts. Accordingly, a theory of truth should "deal chiefly" with the concept of a true thought. For it is on this concept that the constellation of secondary truth concepts depends; it is in virtue of their relation to this concept that sentences may be called true secondarily.[2]

Defects in the semantic conception of truth become evident as soon as one sees truth for sentences as dependent upon the central concept of a true thought. Its most glaring fault is that it completely by-passes the primary concept of a true thought, which is that in virtue of which the indefinitely many semantical truth concepts qualify as truth concepts at all. Doing so, it abandons the possibility of explaining why they are all called truth concepts. Matters are worsened for Tarski's semantic conception by its being stated in terms of the theory of reference rather than the theory of meaning. It attempts to define a sentence's truth in terms of relations among the "references" of its primitive predicates and names. But if a sentence is true because of the truth of the thought it expresses, then the "references" of the sentence's predicates would be only indirectly related to the sentence's truth; for the "references" of the predicates could not determine which proposition a sentence expresses. What the predicates *express* is what is relevant to the thought expressed and, in turn, to the sentence's truth. Furthermore, predicates do not refer to anything in the first place; they only express. Tarski's semantic conception of truth thus has the added trouble of resting on a questionable theory of the fundamental relations between words and things. (See §23 and §38 for an extended critique of referential semantics.) A final problem in Tarski's theory of truth is that it is framed within set theory. But set theory is an artifice without ground in our naturalistic ontology or natural logic and without pragmatic justification either. (See chapter 5 for a critique of set theory.) The theory of qualities and concepts is the proper theory within which to frame a theory of truth.

As Aristotle says, 'everywhere science deals chiefly with that which is primary, and on which the other things depend, and in virtue of which they get their names.' In view of the foregoing discussion, one may conclude that the chief task of a theory of truth is to define the concept of a true thought, for this is the central concept upon which the other truth concepts depend and in virtue of which they are called truth concepts. At the outset of his search for a commonsense theory of truth, Bertrand Russell gives the following reasonable advice:

> ...[We] have to seek a theory of truth which (1) allows truth to have an opposite, namely falsehood, (2) makes truth a property of beliefs, but (3) makes it a property wholly dependent upon the relation of the beliefs to outside things. (p. 123, *The Problems of Philosophy*)

Thus according to Russell, truth is a property of beliefs (i.e., thoughts) that depends upon a relation with things outside, i.e., upon 'a correspondence of thought with something outside thought' (p. 121, *Problems*). And this view is, to the naive eye, a virtual truism. Even those generally unsympathetic to this account of truth acknowledge its privileged position:

> There can be no denying the attractiveness of this view: it seems to be just right. It struck the great philosophers who first considered the problem of truth—viz., Plato and Aristotle—as so obviously the correct one that the question of possible alternatives to it never occurred to them. And certainly if there were such a thing as the common-sense view of truth, it would be the correspondence theory. Common-sense views of this sort may all, in the end, be correct, once they are properly understood; and to call them "common-sense views" is to claim that at the outset they appear to be straightforwardly and undeniably correct. But between the outset and the end (when they are at last "properly understood")—that is to say when they are in the hands of the philosophers—they inevitably run into tough sailing. (p. 4, George Pitcher, *Truth*)

Now if a commonsense view can be made fully clear and precise and if at the same time it can be economically integrated into a larger body of accepted theory, then it is to be preferred over views that clash with it. From Plato and Aristotle down to Russell and Tarski, the correspondence theory of truth has been almost universally acknowledged as the commonsense view. Indeed, the only cogent objection to it has been that it has defied clear and precise formulation. To be sure, a thought is true if and only if it corresponds to a condition in the world, i.e., to a condition that obtains. But what is a thought? What is a condition? What is it for a thought to correspond to a condition? And what is it for a condition to obtain? These are questions that modern correspondence theorists have attempted in vain to answer. The reason for their failure is that they, like nearly all modern philosophers, have been under the spell of representationalism. Their representationalism has blinded them to the fact that the classical correspondence theory is based on realism. The only way to give the correspondence theory a clear and precise formulation is within a realistic framework such as that embraced by Plato and Aristotle.

This is where the theory of qualities and concepts comes in. For, as shown in §42, it harks back to the realism of Plato and Aristotle. Indeed, within the framework provided by this logical theory I have already been able to give definitions of the concepts of a thought and a condition and also of the correspondence relation itself.

Therefore, it remains only to define what it is for a condition to obtain. However, this may be easily done, e.g., as follows:

> x obtains *iff*$_{\text{df}}$ for some property which has an instance, x is just the condition that this property does have an instance.

In symbols, x obtains *iff*$_{\text{df}}$ $(\exists y)((\exists z)z \,\Delta\, y \;\&\; x = |(\exists z)z \,\Delta\, y|^{y})$.[3] Using these definitions, I then define the single central concept of truth:

> x is true *iff*$_{\text{df}}$ x corresponds to a condition that obtains.[4]

What about the semantical and intentional paradoxes? In the ramified type theories proposed by Russell and by Church these paradoxes necessitate an infinite hierarchy of non-equivalent truth concepts for thoughts. However, in the setting of the theory of qualities and concepts such a hierarchy of truth concepts is not needed, for the paradoxes can instead be diagnosed and avoided by means related to those described in §26.[5] Thus, in line with the theory of Plato and Aristotle we have a definition of a single, central concept of truth. Since this definition can be written out entirely in the canonical logical language $\mathscr{L}$ with Δ, here is a clear and precise expression of the classical correspondence theory of truth.

46. Necessity

> ... [T]he terms of *efficacy*, *agency*, *power*, *force*, *energy*, *necessity*, *connexion*, and *productive quality*, are all nearly synonimous; and therefore 'tis an absurdity to employ any of them in defining the rest.
>
> ...[W]hen we speak of a necessary connexion betwixt objects, and suppose, that this connexion depends upon an efficacy or energy, with which these objects are endow'd; in all these expressions, *so apply'd*, we have really no distinct meaning, and make use only of common words, without any clear and determinate ideas. ('Of the idea of necessary connexion', *A Treatise of Human Nature*)

David Hume thus called into question the existence of the concept of a necessary connection. In recent years W. V. O. Quine has called into question the existence of the concept of analyticity. These doubts might seem unrelated. However, like Carnap and many logical positivists, Quine has made a practice of writing as though he thinks analyticity and necessity are the same. In this way Quine may be heard as a contemporary echo of Hume. In fact, the doubts of Hume and Quine have substantially the same origin, being founded in each case on the same sort of argument about definability. The argument goes as follows: if it were to exist, the concept of necessity (analyticity) ought to have a non-circular definition, and yet all the candidate definitions appear upon analysis to be circular. In this section the Humean doubt about the concept of necessity will be met head-on by means of non-circular definitions of necessity and necessary connection. In the subsequent section the Quinean doubt about the concept of analyticity will be met. This two-step strategy is needed, for despite the attitudes of Carnap and Quine, necessity and analyticity have significantly different definitions. It should go without saying that what makes my definitions possible is the theory of qualities and concepts, a logical theory that in relevant respects is much stronger than either Hume's psychology of impressions and ideas or Quine's set-theoretic materialism. Before getting to my definitions of necessity and necessary connection, however, I should like to take a moment to comment on Carnap's popular possible-worlds approach to the problem of necessity. This alternate approach, whose origins may be traced to the writings of Leibniz, has been at once hailed as a worthwhile formal tool and condemned as circular. It would be good, therefore, to get straight on this issue.

In my view both of these judgments of the possible-worlds approach are sound. The reason for this is that there are two quite distinct uses to which this approach is put. First, it is used in formal semantics to characterize validity for certain languages containing necessity and possibility operators. In this application the possible-worlds approach is free of circularity, for the construction is just a part of set theory; in particular, it is a part of model theory. To be sure, certain set-theoretic objects are sometimes spoken of as possible worlds, possible individuals, etc. However, such talk is heuristic in character. In the formal statement of the theory all such talk disappears, and only the vocabulary of first-order extensional set theory remains. Now this application of the possible-worlds approach has been successful within limits. A critic, once he understands the nature of the project, must acknowledge that the class of valid sentences in certain formal languages can be characterized within set theory by means of the possible-worlds technique.

There is, however, another use to which the possible-worlds approach is put; namely, it is used to define necessity itself:

x is necessary *iff*$_{\text{df}}$ x is true in all possible worlds.

Here, the talk of possible worlds is not a mere heuristic, eliminable in favor of set-theoretic talk. On the contrary, 'is a possible world' and 'is true in' are actually primitive constants in the theory. But what is a possible world? And what is it to be true in one? We are as much in the dark on these questions as we are about the nature of necessity. Indeed, the circularity of the possible-worlds definition of necessity can be made explicit. For given the theory of qualities and concepts and the theory of ordinary aggregates (§27)—two theories for which there is independent justification—the primitive terms of the possible-worlds definition of necessity can be defined in terms of necessity. These definitions might go as follows:

> w is a world iff_{df} for every proposition x, x is in w or the negation of x is in w.
>
> x is-true-in w iff_{df} x is a proposition and x is in w.
>
> w is a possible world iff_{df} w is a world and no necessarily false proposition is-true-in w.[6]

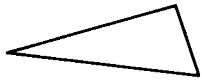
Thus, a "world" is a maximal aggregate of propositions, and a possible world is just one that does not contain necessary falsehoods. Notice, however, that the circularity just exposed in the possible-worlds definition of necessity is precisely the sort that Hume and Quine would criticize. And in this instance at least the criticism seems justified. The conclusion, therefore, is that the possible-worlds approach is not of use in solving the classical problem of necessity.

Let us try another approach. Thoughts are very finely differentiated things; in fact, there is no limit to the number of thoughts that are necessarily true. The following examples will help to indicate the variety there is among the thoughts that are necessary. The thoughts that triangles are triangles, that triangles are three sided, that $5 + 7 = 12$, that the continuum hypothesis is independent of the axioms of Zermelo-Fraenkel set theory, that colors are incompatible, that aesthetic qualities are supervenient: these thoughts are distinct from one another, and they are all necessary. If one is to solve the problem of necessity, one must frame the definition so that every such thought comes out as necessary. Now, recall that a thought is true if and only if it corresponds to a condition that obtains. By analogy, a thought is necessary if and only if it corresponds to a condition that must obtain, i.e., to a condition that is necessary. But I have already given a definition of what it is for a thought to correspond to a condition. Thus, the problem of defining what it is for a thought to be necessary will have been solved if one can define what it is for a condition to be necessary.

So far so good. But now let me consider the matter of necessary connection. Humeans have been unable to make progress on this problem largely as a result of their nominalism and ontological extensionalism. Let me explain. Suppose that there is a necessary connection x between two things y and z. The Humean, being a nominalist, first looks to the individual pair of objects y and z in order to discover what it is for them to have a necessary connection. And, of course, he must fail there. For, as Russell makes clear, relations are located nowhere, and hence, they cannot be discovered in the objects that they relate. Next, the Humean, being an ontological extensionalist, looks to various representative samples of pairs of objects resembling y and z. And, of course, he must fail here too. For connections are intensional and, therefore, can never be adequately characterized by means of samples of their instances, no matter how complete the samples.

The underlying source of the Humean's problem is that he allows himself virtually no *logic*. Now the only reasons that the Humean has for this frugal practice are questionable epistemological ones. I say questionable, for beings possessing so little logical facility could not be deemed rational. We certainly are not beings of that sort. With an appropriate logic, one with which we have a native facility, the problem which only baffles the Humean submits to solution. What I have in mind is the logic for qualities and concepts. Within this logic one is able to define the key logical concept of a connection. The definition is given entirely in terms of the fundamental logical operations. Connections are the special relations from which thoughts can be built up by means of these fundamental logical operations. Given this, if one can also define what it is for a condition to be necessary, then the definition of necessary connection is immediate: x is a necessary connection between y and z if and only if the condition that x is a connection between y and z is a necessary condition. Thus, one has only to consider the matter of necessary condition.

Recall that conditions conform to the first traditional conception of intensional entities, conception 1. Conditions thus are identical if and only if they are necessarily equivalent. Consider, for example, a condition that involves the little black object below:

(1) 

The condition that (1) is three-angled, the condition that (1) is three-sided, the condition that (1) is triangular and such that $5 + 7 = 12$, the condition that (1) is triangular and such that the continuum hypothesis is independent of the axioms of Zermelo-Fraenkel set theory, etc.: these conditions are all the same condition. They are all identical to the same condition right here in the world; it is a condition that I am observing right now. True enough, there is no limit to the number of necessarily equivalent ways I can think about (1) and its shape. However, here in the world there is only one condition to which all these distinct thoughts correspond. Now the same thing goes for necessary conditions. Consider, e.g., the necessary condition that triangles are triangles. What condition is this? It is a condition that must obtain no matter what. It is a way the world must be, come what may. However, there is one and only one way the world must be: it is the way the world necessarily is. There is one and only one condition that must obtain: it is the necessary condition of the world. Thus, the condition that triangles are triangles is a condition in the world that coincides with, e.g., the condition that triangles are three-sided; the condition that $5 + 7 = 12$; the condition that colors are incompatible; the condition that aesthetic qualities are supervenient; etc. These are all the same condition in the world. To be sure, there is no limit to the number of ways one can think about how the world must be; there is no limit to the number of necessary contents of mind (i.e., necessary thoughts) one might have. But they all correspond to the same condition in the world, the condition that a thing is what it is.

Thus, I arrive at the following three definitions:

x is a necessary condition iff_{df} $x =$ the condition that $x = x$.[7]

x is a necessary thought iff_{df} x corresponds to a necessary condition.

x is a necessary connection between y and z *iff*$_{\text{df}}$ the condition that x is a connection between y and z is a necessary condition.

And in the same vein:

$x \mathrel{\dot{\approx}}_{\text{N}} y$ *iff*$_{\text{df}}$ the condition that x is equivalent to y is necessary, and x and y are the same degree.

These definitions can all be written out fully within the purely logical language $\mathscr{L}$ with Δ.

According to this analysis, then, necessity is neither a naturalistic nor an empirical nor a mysterious intuitive concept. It is a logical concept, and a fairly simple one at that. To define necessity, one must only appeal to the fundamental logical operations by means of which conditions and thoughts are formed. Intuitively, the analysis works because these logical operations, together with the genuine qualities and connections upon which they ultimately operate, are the things that determine what is necessary.

It is natural at this point to wonder what the relationship is between necessity, as I have analysed it, and the special kind of necessity known as *logical necessity*. The answer is a truism: logical necessity is that species of necessity having to do with *logic*. Specifically, a thought is logically necessary if and only if it is necessary by virtue of logic alone. It would seem that not all necessary thoughts are logically necessary. For example, it would seem that some are necessary by virtue of metaphysics, and perhaps others are necessary by virtue of causal law. In the next section I will try to define this special concept of necessity. For the present, however, the important point is that logical necessity and metaphysical necessity (and perhaps necessity arising from causal law) are species of the general concept of necessity that was analysed here.

I will close this section with a remark about my Russellian semantic method (defended in §38). Notice that all algebraic model structures $\mathscr{M}$ contain an element $\mathscr{K}$ which, as I indicated in §14, might be thought of as determining various alternate or possible extension functions for the universe of discourse $\mathscr{D}$. Some people might see $\mathscr{K}$ as a vestige of the possible-worlds semantic method and conclude on that basis that my Russellian semantic method must, after all, appeal to possible worlds at least vestigially.[8] However, this would be an error, as I will now explain.

I mentioned in §28 that all the metatheory done in this work should finally be understood, not as part of set theory, but rather as part of a theory of intensional entities. This comment applies to my Russellian semantics: it should be understood as a construction within a general background theory of qualities and concepts. (For example, when one gets around to constructing a serious semantics for fragments of natural language, the universe of discourse will contain intensional entities of the sort provided by the background theory.) Now this background theory will be stated in a first-order extensional language akin to the canonical language $\mathscr{L}$ in which one can express the predication relation and the fundamental logical operations. However, this logical machinery renders the use of algebraic model structures unnecessary. In order to do a Russellian semantics for a fragment of natural language, one need only specify a Russellian interpretation $\mathscr{I}$ and then give a recursive definition of the Russellian meaning function $M_{\mathscr{I}}$. This can all be done straightaway without algebraic model structures (and without $\mathscr{K}$s); one simply uses the special logical machinery provided by the background theory. This, then, is my final proposal for semantics; it makes no allusion to possible worlds.

But where does this procedure leave us on the matters of necessity and possibility? How is one to do the semantics for natural-language talk of necessity and possibility? The answer is that one should do the Russellian semantics for this fragment of natural language just as one would any other fragment. Of course, in this fragment there are expressions whose ordinary meanings are the properties necessity and possibility. Therefore, one will want to choose the interpretation $\mathscr{I}$ so that the relevant expressions are mapped onto these properties. No problems arise here, however, for necessity and possibility are, as I have shown, definable within the theory of qualities and concepts. This is all that is required to obtain a perfectly adequate semantics for the language of necessity and possibility, and it is free of all vestiges of possible worlds.[9]

47. Analyticity

In all judgments in which the relation of a subject to the predicate is thought (I take into consideration affirmative judgments only, the subsequent application to negative judgments being easily made), this relation is possible in two different ways. Either the predicate B belongs to the subject A, as something which is (covertly) contained in this concept A; or

B lies outside the concept *A*, although it does indeed stand in connection with it. In the one case I entitle the judgment analytic, in the other synthetic. (Immanuel Kant, *Critique of Pure Reason*)

By use of this metaphorical language Kant introduced his famous concept of analyticity, a concept that would play a central role in his own system of philosophy and in much of the subsequent philosophy in the European tradition. The concept was by no means original with Kant. Closely related concepts had played a role in the thought of other modern philosophers: Locke's trifling propositions, Leibniz's identical propositions, Hume's relations of ideas. The problem of analyticity confronting philosophers today is that of giving a precise, non-circular definition of the concept, a definition which facilitates answers to the fundamental epistemological questions that troubled Kant originally.

The Quinean position is that the problem has no solution:

But for all its *a priori* reasonableness, 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. (p. 37, W. V. O. Quine, 'Two Dogmas of Empiricism')

Now some might think that a solution to the Humean problem of necessity can double as a solution to the problem of analyticity. Indeed, following in the footsteps of Carnap, Quine often writes as though the concepts of analyticity and necessity were interchangeable. I will argue that they are not, however, and if I am right we must look elsewhere for a definition of the Kantian concept.

In 'Notes on the Theory of Reference' Quine paints the following pessimistic picture:

In Tarski's technical construction ... we have an explicit general routine for defining truth-in-*L* for individual languages which conform to a certain standard pattern and are well specified in point of vocabulary. We have indeed no similar single definition of 'true-in-*L*' for variable '*L*'; but what we do have suffices to endow 'true-in-*L*', even for variable '*L*', with a high enough degree of intelligibility so that we are not likely to be averse to using the idiom.

See how unfavorably the notion of analyticity-in-*L*, characteristic of the theory of meaning, compares with that of truth-in-*L*. [We have no] ... systematic routine for constructing definitions of 'analytic-in-*L*', even for the various individual choices of *L*; definition of 'analytic-in-*L*' for each *L* has seemed rather to be a project unto itself. The most evident principle of unification, linking analyticity-in-L for one choice of L with analyticity-in-L for another choice of L, is the joint use of the syllables 'analytic'. (p. 138)

But what Quine says here does not hold up. There is a general routine for defining analyticity-in-L, one that is fully comparable to Tarski's routine for defining truth-in-L. In order to arrive at this routine for defining analyticity-in-L, one must only combine the formal algebraic semantics of §14—which leads to a definition of validity-in-L—and the Russellian semantics of §38—which yields a definition of meaning-in-L. The resulting routine yields a neat little definition of analyticity-in-L:

> A is analytic-in-L relative to interpretation $\mathscr{I}$ and algebraic model structure $\mathscr{M}$ *iff*$_{\mathrm{df}}$ there is a valid L-formula B such that, for some interpretation $\mathscr{I}'$, the meaning of B relative to $\mathscr{I}'$ and $\mathscr{M}$ is the same as the meaning of A relative to $\mathscr{I}$ and $\mathscr{M}$.

i.e.,

$$\mathrm{An}_{\mathscr{I}\mathscr{M}}(A) \ \textit{iff}_{\mathrm{df}}\ (\exists B)(\models B\ \&\ (\exists \mathscr{I}')\mathrm{M}_{\mathscr{I}\mathscr{M}}(A) = \mathrm{M}_{\mathscr{I}'\mathscr{M}}(B)).$$ [10]

Though the doubt raised by Quine in 'Notes on the Theory of Reference' can be resolved in this way, Quine would not feel that the original problem of defining analyticity had been solved. For to employ the above routine in the case of a particular spoken language L, one must first know the correct interpretation $\mathscr{I}$ for the primitive predicates and names in L. Yet Quine thinks there are insurmountable barriers to such empirical semantic knowledge.[11] Now although I believe such skepticism can be overcome, this epistemological controversy is a side issue. As I will show later, the concept of an analytic proposition is directly definable without recourse to any semantical concepts. Thus, the solution to the original problem of defining analyticity does not ride on the possibility of empirical semantic knowledge; it depends only on whether one has an adequate theory of propositions.

Quine links the problem of analyticity to a problem in the theory of empirical knowledge because he embraces what may be called the logical positivist conception of analyticity. This conception is distinctive in two ways: (1) it treats interpreted sentences as the primary bearers of analyticity, and (2) it treats an interpreted sentence as analytic if and only if the sentence is alike in meaning to a valid sentence, where validity is understood in Tarski's model-theoretic way. Quite apart from Quine's skepticism about empirical semantic knowledge, there are reasons not to embrace the logical positivist conception. First, I have shown (§45) that propositions, not sentences, are the primary bearers of truth (falsehood); sentences are true (false) only secondarily through their meanings. It would seem by analogy that propositions should be identified as the primary bearers of analyticity and that sentences should be counted as analytic (synthetic, contradictory) only secondarily through their meanings. Secondly, there are, I will argue, difficulties in the Tarskian model-theoretic conception of validity. Before I consider that issue, however, it will be helpful to explore the implications of the thesis that propositions are the primary bearers of analyticity.

Since the time Kant introduced his concept of analyticity there have been only a few attempts to find a clear and precise definition for it. One attempt is found in Carnap's *Meaning and Necessity*. In this work, however, Carnap in effect identifies analyticity with necessity: '"*L-true*" is meant as an explicatum for what Leibniz called necessary truth and Kant analytic truth' (p. 8). (Note that Carnap defines 'L-true' in terms of his state descriptions, which he indicates (p. 9) are intended to '...represent Leibniz' possible worlds or Wittgenstein's possible states of affairs'.) Continuing the Carnapian approach, David Lewis also in effect identifies analyticity with necessity. (P. 174, *Convention*. Note that Lewis identifies necessity with truth-in-all-possible-worlds.) In my view, this approach to analyticity is not what is wanted. Kant held that whatever is knowable *a priori* is necessary and, hence, that whatever is synthetic *a priori* is a synthetic necessity; the greater part of the *Critique of Pure Reason* is devoted to the study of these synthetic necessities. However, given the possible-worlds definition, it follows by a one-step inference that synthetic necessities could not exist.[12] Kant might have been mistaken about the existence of synthetic necessities, but if so, the mistake was a deep theoretical one. We do not want to undercut trivially Kant's philosophy and the tradition surrounding it simply by the way we define analyticity. Rather, we want to sharpen Kant's informal, metaphorical definition in a way that enhances the investigation of the existence of synthetic necessities.

The problem with the possible-worlds definition of analyticity stems from its reliance on the traditional conception of intensions according to which they are identical if and only if they are necessarily equivalent—that is, its reliance on conception 1. Since on this conception there is only one necessity and since all analyticities are necessary, it follows all too quickly that there can be no synthetic necessity. The way out of this problem is to give due weight to the fact that in the Kantian scheme *judgements* are primary bearers of analyticity. 'Judgement', like 'thought', is ambiguous. It can be used to mean a kind of intentional act and can also be used to mean a type of proposition that is the characteristic object of that kind of intentional act. In turn, there are two uses of 'analytic', one for intentional acts of judging and one for the propositional objects of those intentional acts. These two uses of 'analytic' are related by the following elementary equivalence: the intentional act of judging the proposition x is analytic if and only if the proposition x is analytic. Since the relation between the two uses is so direct, one need not be concerned here with the issue of whether the intentional act or the propositional object should be taken as the primary bearer of analyticity,[13] one should feel free to adopt either alternative. And since the latter alternative leads to a simpler treatment, I will adopt it as a matter of convenience. This practice is consonant with much, though not all, of Kant's own usage. So what type of 0-ary intensions are typically the objects of these intentional acts (judgements, thoughts, denials, hunches, recollections, etc.)? They are, of course, the type that fall under traditional conception 2; that is to say, they are thoughts. Thoughts may thus be identified as the primary bearers of analyticity.[14]

Since Kant's time the theory of logical form has undergone significant change. We entertain a much richer system of formal classification of thoughts than Kant did. Kant mentions only two categories of thoughts in his definition of analyticity, affirmative and negative subject/predicate thoughts, and he explicitly defines analyticity only for affirmative subject/predicate thoughts. But he does say, 'I take into consideration affirmative judgments only, the subsequent application to negative judgments being easily made'. The implication is that there are negative analytic thoughts, too, and that the general notion of analyticity is to be obtained by appropriately adapting the circumscribed definition to this further category of thought. (So, e.g., if it is analytic that all A are B, then it is also analytic that all things that are not B are not A; etc.) In view of this, it is natural to ask what Kant would say about the still further categories of thoughts entertained by our logical theory.[15]

Presumably, he would adopt the same attitude toward these further categories as he did toward the category of elementary negative thoughts. (If not, what good reason could he have for drawing the line here, just after admitting negative analytic thoughts?[16]) In that event, the general Kantian concept of analyticity must be obtained, not by piecemeal extensions of the original circumscribed definition, but rather by generalization on the essential underlying feature he was trying to get at in the original.

When one performs this kind of generalization, one arrives at more or less the following. Analytic thoughts are those that must be true by virtue of logic alone; their particular non-logical content is immaterial. But notice, thoughts that must be true are necessary, and thoughts that are necessary by virtue of logic alone, independently of their non-logical content, are necessary logically. Now, for Kant, not all necessary thoughts are necessary logically. Some are necessary by virtue of their non-logical content. For Kant, the necessities that are logical are analytic, and the necessities that are non-logical are synthetic.

When we say of a thought that it must be true by virtue of logic alone independently of its non-logical content, what we mean is that the thought is one that must be true because of its logical form and that its non-logical content is immaterial. However, thoughts that must be true by virtue of their logical form are none other than those that are valid. Thus I arrive at the following conclusion. An analytic thought is just a thought that is valid, where a valid thought is one whose necessity is logical rather than non-logical in nature. The problems of defining analyticity, validity, and logical necessity are consequently one and the same. Now a thought is made necessary by its logical form (independently of its non-logical content) if and only if any proposition having the same logical form (though perhaps a different non-logical content) is necessary. So the problem of defining validity (analyticity, logical necessity) turns on the question of what the logical form of a thought is.

There are two opposing views on this question. According to the first, the logical form of a thought is simply the abstract shape (form) of its complete thought-building tree, i.e., the tree determined by the inverses of the fundamental thought-building operations. Besides these fundamental logical operations, the identity of the other nodes (i.e., the content) in the tree is immaterial. The second view of logical form is just like the first except that the identity of all purely logical nodes is counted in. That is, the logical form of a thought is the form of its complete thought-building tree when the purely logical content is held constant.

The second view is, I maintain, clearly the right one. For it is only on this view that elementary validities involving, say, identity and necessary equivalence (e.g., $[(\forall x)x = x]$ and $[(\forall x)x \approx_N x]$) qualify as valid. However, if the purely logical relations of identity and necessary equivalence are counted in, then so must the purely logical relation predication. The predication relation is no less a logical relation than are identity and necessary equivalence. In fact, I have shown that identity and necessary equivalence are definable in terms of the predication relation (together with the fundamental logical operations).[17] And just as there are highly intuitive elementary validities involving identity and necessary equivalence, there are highly intuitive elementary validities involving the predication relation (e.g., $[(\forall x)(x \text{ lives} \equiv x \text{ is living})]$).[18] In this matter, the predication relation is for Kant the paradigm of a purely logical relation; it is central throughout his considerations of what it takes for a thought to be analytic. Thus, I conclude that the logical form of a thought is determined by all the purely logical elements that show up in its analysis under the inverses of the thought-building operations; in particular, the predication relation is to be counted as one of these purely logical elements.

With this conclusion in hand, let us consider the Tarskian model-theoretic account of validity, which underlies the logical positivist conception of analyticity. The goal of the Tarskian account is to define in exclusively set-theoretic terms what it takes for a sentence in a given language to be valid. Everyone must admit that the Tarskian account of validity achieves its goal for a number of interesting languages and that this is of considerable value. Nevertheless, a significant obstacle seems to stand in the way of a general Tarskian account of validity.[19] This obstacle is tied to the fact that the logical form of a thought is determined in part by occurrences in its analysis of the purely logical relation *predication*.[20] Let me explain. On the usual Tarskian account, the validity of a sentence is just truth-in-*all*-possible-models. However, when a predicate (e.g., $=$, $\approx_N$, Δ) is singled out as a distinguished logical predicate, the Tarskian must appropriately narrow the class of models if he is to get the right result. For example, when $=$ is singled out as a distinguished logical predicate, the class of models is narrowed down so as to include only those models $\langle \mathscr{D}, \mathscr{R} \rangle$ in which the "reference" of the predicate $=$ is just the extensional identity relation on $\mathscr{D}$, i.e., $\mathscr{R}(=) = \{xy \in \mathscr{D} : x = y\}$. In §17 I showed how a comparable narrowing of the admissible models can be attained for the distinguished logical predicate $\approx_{\mathrm{N}}$. The problem facing the Tarskian account of validity is to do the same thing for the distinguished logical predicate Δ, which on the intended interpretation expresses the predication relation. Gödel's first incompleteness theorem shows that there is no syntactical solution to the problem. And in §26 I showed that a model-theoretic version of Russell's paradox results if one attempts to solve the problem by naively requiring that $\mathscr{G}(\bar{\Delta}) = \{xy : x \in \mathscr{G}(y)\}$. The only way I know to guarantee a correct model-theoretic account is to *use* a Δ-predicate in the metatheory itself and, thereby, explicitly to transgress the set-theoretic limits imposed by Tarski. But if I am right about this, why bother to go to all the trouble to give a model-theoretic account? It is far more simple and natural to dismantle the model-theoretic superstructure and to give the definition of validity within the object-theory itself, i.e., within the full theory of qualities and concepts with the predication relation. In so doing, one should take the theory at face value on its intended standard interpretation. There is no shame in this strategy, for set theory—not the theory of qualities and concepts—is the theory that one should be happy to do without. By the stage when one is framing a general definition of validity, the Tarskian model-theoretic approach has outlived its usefulness.

The best strategy, therefore, is to define straightaway what it is for a thought to be valid. Then and only then can a general definition of validity for sentences be given: a sentence is valid in (a fragment of) a language *iff*$_{\mathrm{df}}$ it expresses a valid thought in (the fragment of) the language. Thus, there can be an adequate general definition of validity for sentences if and only if there is an adequate definition of the Kantian concept of analyticity. Ironically, by attacking the possibility of an adequate definition of analyticity, philosophers unwittingly attack the possibility of an adequate general definition of validity for sentences.

Let me turn now to the definition of analyticity. A thought, I have said, is analytic if and only if every thought having the same logical form is necessary. What I must do now is to turn this informal definition into a formal one. Assume for a moment that we know what a *purely logical object* is (i.e., the sort of object that is a purely logical node in a thought-building tree). Let t and t' be two complete thought-building trees (i.e., thought-building trees that cannot be analysed further by means of the inverses of the thought-building operations). Suppose that t and t' can be obtained from one another by making replacements among their non-logical nodes in accordance with the following rule: for all objects v and v', if v is a node in t and v' is the associated node in t', then v is found in t at all and only those places where v' is found in t'. Let u be the thought that has tree t, and let u' be the thought that has tree t'. In this case u and u' are defined to have the same logical form. Then, analyticity may be defined in the intuitive way:

> A thought is analytic iff_{df} every thought having the same logical form is necessary.

Suppose that a predicate for purely logical objects is adjoined to the purely logical language $\mathscr{L}$ with Δ. Then, given that there are not complete thought-building trees having infinitely long branches, one can write out this definition of analyticity in a Zermelo-style theory for Δ.

It would be nice to have a definition of what a purely logical object is. I will suggest one. It should be noted, though, that the correctness of the definition of analyticity does not ride on the correctness of this definition, for if worst comes to worst, the notion of a purely logical object could be taken as undefined. Thus far I have seemed successful in defining a variety of purely logical concepts in terms of the predication relation and the fundamental condition-building and thought-building operations. This prompts the conjecture that all and only purely logical objects can be built up from the fundamental logical relation of predication by means of these two types of fundamental logical operations. This conjecture is in the spirit of this work. And given the role predication and these fundamental logical operations play in the world and in thought, I do not see any clear-cut counterexamples. However, the conjecture is distinctly philosophical and has no positive proof. In any case, if it is correct, then the conjecture can be converted into a definition of the concept of a purely logical object.[21] And, if correct, this definition suggests, in turn, a definition of the concept of a *logical constant*: a logical constant is an expression whose meaning is a purely logical object.

From the definitions in this chapter emerges an intuitive picture of three central logical concepts that have loomed large in modern philosophy: truth, necessity, and analyticity. A thought is true if and only if it corresponds to a condition that obtains; a thought is necessary if and only if it corresponds to a condition that must obtain, and a thought is analytic if and only if every thought having the same logical form corresponds to a condition that must obtain. I have also argued that validity and analyticity are one and the same. If I am right, then the definition of analyticity doubles as a definition of validity:

> A thought is valid $\mathit{iff}_{\mathrm{df}}$ every thought having the same logical form corresponds to a condition that must obtain.[22]

I will bring the chapter to a close with some remarks on the conception of logic that emerges from this definition of validity. This is called for since logic, by definition, is concerned with validity and valid thinking.

According to Aristotle's conception, logic is primarily a *tool*, or *organon*, for valid thinking. True, Aristotle believed that logic inevitably touches questions about the basic components of and structure of thoughts and that, in this, logic overlaps certain fundamental parts of metaphysics. But this substantive dimension was viewed by Aristotle as incidental. At the onset of modern philosophy Francis Bacon sought to revise the Aristotelian conception of logic by expanding its scope so as to include certain forms of inductive reasoning. It is doubtful that the expanded Baconian conception is warranted; in any event it did not call into question the underlying Aristotelian view that logic is a tool for valid thinking. Perhaps the first truly major alteration in the Aristolelian conception came with Frege. Initially, Frege too approached logic as a tool for valid thinking. However, in time he realized that an adequate formulation of logic required positing a wide range of purely logical objects. Indeed, (numerical) mathematics turned out to be nothing but a science that deals with a special kind of purely logical object. Thus it was that logic came to have a legitimate subject matter of its own, a subject matter that was not merely incidental as in the case of Aristotle's organon.[23] Nevertheless, someone could still sustain the belief that all the valid thoughts, and, hence, all the valid ways of thinking, could be captured by a well constructed organon. And thus, someone could still hold that logic, at least as defined by its purpose, is a tool for valid thinking. In this, the Aristotelian conception still seemed viable. However, this situation was brought to an abrupt end by Gödel. Given Gödel's first incompleteness theorem and given the logicist thesis that all truths of (numerical) mathematics are validities, logic had to be viewed as a full-fledged, evolving science in its own right. Its laws, i.e., the valid thoughts, could never all be captured by a tool. Valid thoughts—and, in turn, valid ways of thinking—would always be left out no matter how well the job is done. These valid thoughts would have to be discovered by some other means than the application of a tool. To be sure, among the by-products of logic there are numerous tools to aid valid thinking. But logic finally had to be considered a science primarily and a tool only secondarily through these by-products. And so it was that the Aristotelian conception of logic became untenable.

Few people today would accept this historical sketch. Yet I think that it is close to the truth. Two things are wrong with it, however. First, Frege's purely logical analysis of number broke down in some of its details. Secondly, by the time Gödel proved his result, it was hardly taken for granted that the truths of mathematics are part of logic; indeed, Gödel himself rejected the logicist thesis. Thus, the above sketch does not hold up as history. Despite this, the conclusions that logic has its very own ontology and that the Aristotelian organon conception of logic is untenable can be won by virtually the same route. One need only fill in the two gaps left by history. In effect, I have tried to do this. First, I gave a neo-Fregean analysis of number which, it has been argued, is free of the flaws present in Frege's original analysis. Secondly, I gave a defense of the logicist thesis that all truths of (numerical) mathematics are valid. This thesis follows immediately from the neo-Fregean analysis of number plus the proposed analysis of validity.

Logic, therefore, is not a tool; it is an open-ended, evolving science having an ontology of its own. Mathematics is but one of its parts, and its full scope is yet to be discovered. Given the definition of validity, every necessary thought whose analysis contains only purely logical objects is valid. So every time we discover a purely logical analysis of a concept previously thought to belong to a discipline outside logic, the acknowledged scope of logic must be expanded accordingly. The purely logical analysis of number is just one case in point, and I submit that there are many more. In this vein, I will in the closing chapter venture into an area that on the face of it might seem to some as foreign to logic as mathematics once did, namely, the area of intentionality, mind, and consciousness. If there is anything in the analyses I will offer, then the conception of logic that emerges is very far indeed from that of Aristotle, for whom logic is primarily a tool; instead, it is more like that of Plato, for whom logic is akin to reason itself.

10

Mind

48. Intentionality

> Every mental phenomenon is characterized by what the scholastics of the Middle Ages called the intentional (and also) mental inexistence of an object, and what we would call, although not in entirely unambiguous terms, the reference to a content, a direction upon an object (by which we are not to understand a reality . . .), or an immanent objectivity. Each one includes something as an object within itself . . . .
>
> This intentional inexistence is exclusively characteristic of mental phenomena. No physical phenomenon manifests anything similar. Consequently, we can define mental phenomena by saying that they are such phenomena as include an object intentionally within themselves. (Franz Brentano, *Psychologie vom empirischen Standpunkt*)

An intentional phenomenon, according to Franz Brentano, is one that makes reference to, is directed upon, or is about other objects, perhaps even objects that do not exist. Intentional phenomena can in this sense be said to 'include an object intentionally within themselves'. Intentionality, then, is that special property of being directed upon something (*Gerichtetsein*). Brentano used this concept of intentionality to formulate a two-part thesis that has come to be known as *the thesis of intentionality*:

(1) All and only mental phenomena are intentional.
(2) No purely physical phenomenon is intentional.

This is to say, (1) intentionality, i.e., the special property of directedness or aboutness, is the mark of the mental, and (2) it sunders the mental from the purely physical. I will return to Brentano's thesis in a while, but at the moment my concern is with the concept of intentionality itself. For though Brentano is to be credited with the modern rediscovery of intentionality, his analysis of it is inadequate. My immediate goal is to define intentionality without appealing to the metaphors of directedness and inexistence.

I will begin by giving a schematic summary of Brentano's theory of judgement as it is reported by Roderick Chisholm.[1] Brentano's theory differs sharply from the propositional/relational theory of judgement that I have been espousing in this work. On Brentano's theory, when one judges that $(\exists x)Ax$, one does not stand in relation to the proposition that $(\exists x)Ax$; nor does one stand in a relation to the concept of being an *A*. Instead, one affirms or accepts *A*s. Likewise, when one judges that $\neg(\exists x)Ax$, one does not stand in a relation to the proposition that $\neg(\exists x)Ax$. Rather, one denies or rejects *A*s. In the same vein, to judge that $(\exists x)(Ax \& Bx)$ is to accept *A*s that are *B*s. To judge that $\neg(\exists x)(Ax \& Bx)$ is to reject *A*s that are *B*s. To judge that $(\exists x)(Ax \& \neg Bx)$ is to accept *A*s that are non-*B*s, and to judge that $(\forall x)(Ax \supset Bx)$ is to reject *A*s that are non-*B*s. In increasingly awkward steps Brentano thus attempts to extend his theory to complex judgements.

Non-propositional/non-relational theories of judgement are not rare. Evidence of them is found in works ranging from Plato's *Sophist* (240D, 260C–263D) and *Theaetetus* (188E–189A) to Russell's *The Problems of Philosophy* (chapter 12) and his introduction to the first edition of *Principia Mathematica* (pp. 43–4). Yet all such theories share a flaw, indeed, the very flaw that spells defeat for adverbial and multiple-operator approaches to intensional logic. (See §6.) Even if by various awkward maneuvers these theories can handle statements concerning particular judgements, they cannot handle general statements concerning judgements. To handle general statements, one must be able to bring the theory within the scope of quantifier logic, and this is precisely what non-relational/non-propositional theories are unable to do in a credible way. Consider, e.g., the following intuitively valid arguments:

Whatever x believes is necessary.
Whatever is necessary is true.
∴ Whatever x believes is true.

Whatever x believes is true.
x believes that *A*.
∴ It is true that *A*.

x believes that *A*.
∴ x believes something.

I argued in §6 that on the canonical syntactic treatment of such arguments 'believes' is represented as a 2-place predicate and 'that *A*', as a singular term. Only then do the arguments submit to a plausible treatment within quantifier logic:

$$\begin{array}{ll} & (\forall y)(B(x,y) \supset Ny) \\ & (\forall y)(Ny \supset Ty) \\ \hline \therefore & (\forall y)(B(x,y) \supset Ty) \end{array}$$

$$\begin{array}{ll} & (\forall y)(B(x,y) \supset Ty) \\ & B(x,[A]) \\ \hline \therefore & T[A] \end{array}$$

$$\begin{array}{ll} & B(x,[A]) \\ \hline \therefore & (\exists y)B(x,y). \end{array}$$

Once this conclusion is reached, however, one is obliged to determine what special values are included in the range of the variable 'y' and to what the singular term 'that *A*' is semantically correlated. I argued in §8 that on the canonical semantical treatment propositions are the entities that fill the bill. I then concluded that, since 'believes' is a 2-place predicate, it expresses a binary relation whose range is made up of propositions. And this conclusion leads directly to the relational/propositional theory of judgement. However, on the non-relational/non-propositional theory of judgement there is no credible way even to express the above intuitively valid arguments, for this theory denies from the start that belief consists in standing in a relation to truth bearers. The non-relational/non-propositional theory thus falters at the earliest possible stage: it collides with logic itself.

Not only does the relational/propositional theory of judgement mesh easily with logical theory, it also makes possible the first step toward clarifying the phenomenon of intentionality, a step that is out of the question for Brentano's theory and the other non-relational/non-propositional theories. Let me explain. It is commonplace to say that thoughts, beliefs, judgements, etc. are about or directed toward other objects. This aboutness or directedness is what Brentano means by the intentionality of thought, belief, judgement, etc. Now each expression in the family 'thought', 'belief', 'judgement', etc. has at least three related uses. Each can be used to mean (1) a kind of intentional act, (2) the propositional object of the intentional act, or (3) a relation holding between persons performing the intentional act and the propositional object of the act. The non-relational/non-propositional theory acknowledges only the first of these three uses, the one for intentional acts. This forces the theory to give its account of intentionality in the inevitably opaque terms of intentional acts, making metaphor and circularity unavoidable. By contrast, the relational/propositional theory acknowledges all three uses and, thus, is free to analyse intentional acts in terms of the associated relations and their propositional objects. The following is an illustration of how easy these analyses can be: x performs the intentional act of thinking that A if and only if x stands in the thinking relation to the thought that A.

However, I have said nothing yet concerning the intentionality of intentional acts, i.e., their directedness or aboutness. How does that arise? The answer is that it arises from the propositional objects, i.e., from the thoughts to which the person stands in the relation thought, belief, judgement, etc. After all, thoughts in the propositional sense are themselves things that are characteristically said to be about other objects; indeed they are often said to be about objects that do not exist. And the same thing holds not just for thoughts but for all complex ideas.[2] (Unlike complex ideas, simple ideas are just qualities, connections, and conditions. As such, they are not said to be about anything. E.g., the color red is not said to be about anything. Qualities simply qualify things; connections simply connect things; and conditions simply obtain. For example, recall the little triangle figure displayed in §40. One might say, e.g., that a set containing that triangle "involves" or "pertains to" the triangle but not that it is "about" the triangle; analogously, though the condition that the little figure is triangular involves or pertains to the triangle, it is not about the triangle. The condition is just there on the page. An intensional entity can be said to be about other things only if it has a complex logical form.[3]) Now a complex idea—whether it be a thought or a complex concept—can be said to be about other things even when the idea is not the object of any intentional act. (There are thoughts about Hamlet that no one has as yet had.) An intentional act, on the other hand, can be said to be about those and only those things that the associated complex idea can be said to be about. These facts strongly suggest that an intentional act can be said to be about other things for one reason only; namely, the intentional act consists in standing in a certain relation to a complex idea that can be said to be about things. This is to say, an intentional act can be said to be about other things only secondarily through the complex idea that is the object of the act. Complex ideas are the objects that can in a primary way be said to be about other things.

But what features do complex ideas have that allow us to say that they are about other things, including even things that do not exist? Using the theory of qualities and concepts, we can answer this question. The answer goes in pragmatic stages, however, since from a linguistic point of view this idiom of aboutness is an extremely context-dependent affair.

The tightest way in which a complex idea might be said to be about an object occurs when the idea is the result of predicating something of the object itself. For example, the thought that Fx might be said to be about the object x since this thought is the result of predicating $[Fy]_y$ of the object x. A second and somewhat weaker way in which a complex idea might be said to be about an object occurs when the idea is descriptive in character, i.e., when the idea can be denoted by an intensional abstract of the form $[A(\imath x)(Fx)]_\alpha$. Here the idea might be said to be about the unique object to which the associated descriptive concept applies, i.e., about the unique object that stands in the Δ-relation to the concept $[Fx]_x$. In this case the object need not even exist: for example, in an appropriate conversational context I might truly say that the thought that the golden mountain does not exist is about the golden mountain.[4] A third and rather weak way in which a complex concept might be said to be about an object occurs when the concept applies to the object (i.e., when the object stands in the Δ-relation to the concept). In this vein, a thought might be said to be about an object if the analysis of the thought contains a concept that applies to the object. Fourthly, there is a very weak way in which a complex concept might be said to be about objects. For any complex concept $[A(\alpha)]_\alpha$, we might in an appropriate conversational context say that $[A(\alpha)]_\alpha$ is about As regardless of whether any As actually exist. Analogously, we might say that a thought is about As simply if a concept $[A(\alpha)]_\alpha$ occurs in the analysis of the thought. For example, we might say that the thought that no witches exist is about witches. Finally, in addition to the above ways in which a complex idea might be said to be about objects, there are myriad intermediate ways that depend on fine points of logical form, antecedently determined interests, mutually held beliefs, etc.

Now every complex idea in some context or other could be said to be about something, even if that something does not exist. Which thing a complex idea is said to be about depends in large part upon the conversational context. In one conversational context, for example, it might be appropriate to assert that the thought that the witch blighted the sheep is about the witch, and yet in another, to deny this on the grounds that there are no witches. It all depends on what is deemed relevant in the context.

According to this analysis, then, the word 'about' has no semantically fixed extension; its extension is pragmatically determined. But we should not expect this to be otherwise, any more than we should expect the word 'relevant' to have a semantically fixed extension. Indeed, 'about' and 'relevant' have much in common: 'What are you talking about?' is very close to 'What is relevant to your conversation?'. The important point about the above analysis is that it successfully identifies all the formal features of complex ideas that determine what in a given context the thought or concept can be said to be about. And it accomplishes this using only purely logical terms provided by the theory of qualities and concepts.

So far I have reached the following conclusions. First, there are independent logical grounds supporting the relational/propositional theory of judgement. Secondly, using the relational/propositional theory, one is able to analyse intentional acts in terms of the associated relations and propositional or conceptual objects. Thirdly, the intentionality (i.e., the directedness or aboutness) of an intentional act can be accounted for by the fact that the intentional act consists in standing in an appropriate relation to a complex idea—either a thought or a complex concept—which given the right context can always be said to be about other objects, even objects that do not exist. Fourthly, using the theory of qualities and concepts, one can identify all the formal features that are at work in determining what in a given context a complex idea can be said to be about.

Yet the story is not complete, for there is an unsolved problem. Standing in just any relation to a complex idea does not constitute an intentional act. Only certain very distinctive relations will do—relations such as thinking, believing, judging, remembering, perceiving, desiring, deciding, intending, etc. These relations, naturally enough, are called intentional relations. The problem is to give a non-circular definition of what an intentional relation is. If this problem can be solved, then the analysis of intentionality will be complete.

In the English-speaking world interest in intentionality and Brentano's thesis has been generated to a great extent through the efforts of Roderick Chisholm. Many of Chisholm's ideas on intentionality are expressed in his well-known published correspondence with Wilfrid Sellars 'Intentionality and the Mental'. In this correspondence Sellars asserts that every intentional sentence can be analysed into some sentence that uses only non-intentional vocabulary. In opposition, Chisholm echoes the second half of Brentano's thesis of intentionality by asserting that no intentional sentence can be analysed without appeal to further intentional vocabulary. Chisholm, however, also maintains that all intentional sentences have certain purely logical properties that are not shared by any non-intentional sentences. A bit later I will return to the Chisholm/Sellars dispute. At present, it is Chisholm's purely logical criterion for intentional sentences that interests me.

Over the years Chisholm has offered a variety of criteria. His best one appears in his article 'Intentionality' in *The Encyclopedia of Philosophy*. There he attempts to define the narrower concept of a *simple intentional sentence prefix*; he does this in the hope that the wider concept of an intentional sentence can then be defined in terms of this narrower concept. A sentence prefix is any expression that, when prefixed to a sentence, yields a new sentence. A simple sentence prefix is one that contains no meaningful proper part that is a sentence function, where a sentence function is a special kind of expression that, when supplied with a further expression, yields a sentence. Chisholm's definition is this:

> A simple sentence prefix M is intentional if and only if for every sentence A, $M(A)$ is a contingent sentence.

Consider some examples. The simple sentence prefix 'John believes that' qualifies as intentional according to the definition since every sentence (e.g., '$1 + 1 = 3$') is such that the result of prefixing 'John believes that' (e.g., 'John believes that $1 + 1 = 3$') is contingent. By contrast, 'necessarily' does not count as intentional because there are sentences (e.g., '$1 + 1 = 3$') such that the result of prefixing 'necessarily' (e.g., 'necessarily, $1 + 1 = 3$') is not contingent.

Chisholm's ingenious definition seems to be founded on two insights into the nature of intentionality. The first harks back to Brentano's original characterization of intentional phenomena as those that are about objects and perhaps even about objects that do not exist. I have argued that intentional acts consist in standing in intentional relations to thoughts or to complex concepts and that thoughts and complex concepts are the sort of thing that can be said to be about other things, perhaps even objects that do not exist. Now if a relation can hold contingently between an individual and a thought or complex concept, then typically the relation will hold between the individual and the thought or concept independently of whether the thought is true and independently of whether the concept applies to anything. (Factive intentional relations, e.g., knowing, are an exception to this; but since they are definable ultimately in terms of non-factive intentional connections, e.g., believing, Chisholm's insight stands with this qualification.) However, if a thought is not true and if a complex concept does not apply to anything, then in an appropriate context we could say that they are about objects that do not exist. This, then, is the link up between Chisholm's definition and Brentano's original characterization. The second insight upon which Chisholm's definition seems to be founded concerns the concept of a phenomenon. If the condition that a certain relation holds between a certain pair of objects is either necessary or impossible, then the condition cannot be considered to be a phenomenon. The condition would be a phenomenon only if it were contingent. The insight is that fundamental intentional relations are ones that give rise to genuine phenomena and, therefore, must be able to hold contingently at least between certain relata.

Though Chisholm's definition helps us to uncover these two important insights into the nature of intentionality, it encounters a number of difficulties. To begin with, since a sentence prefix is a kind of operator, Chisholm's treatment of intentional language must be classified as a special case of the multiple-operator approach to intentionality. In this, his definition contains vestiges of the non-relational theory of judgement. Not surprisingly, then, general intentional sentences, which spelled defeat for the non-relational theory, also appear to cause trouble for Chisholm's treatment of intentional language. For example, since 'Whatever *x* believes is true' contains the predicate 'believes', which expresses an intentional relation, it would seem to qualify as an intentional sentence. However, it is not clear that Chisholm's treatment helps to show this, for predicates are not sentence prefixes. It would seem better, therefore, simply to bypass Chisholm's linguistic superstructure and instead to define the concept of intentional relation straightaway. This is what I advocate.

To see the other difficulties with Chisholm's definition, one should realize why Chisholm limits the definition to just those sentence prefixes that are simple (i.e., to those sentence prefixes that contain no meaningful proper parts that are sentence functions). For if he did not limit it this way, then numerous easy counter-examples (generated by "meaningful proper parts") would arise. Now if the definition does any more than isolate some accidental feature of, say, English, it should apply to any language that might come to be spoken. But in that case, the very same sort of counterexamples beset the definition in spite of its limitation to simple sentence prefixes. For example, let English* be the language that results when English is supplemented with certain new vocabulary items. In English* let M_1 be a syntactically primitive sentence prefix that is synonymous to the complex sentence prefix:

Socrates is Greek & (______ or not ______).

Now observe that, for any sentence A, $M_1(A)$ is a contingent sentence. Take an example:

Socrates is Greek & ($1 + 1 = 3$ or not $1 + 1 = 3$).

is a contingent sentence. Thus, $M_1(1 + 1 = 3)$ is contingent. In this way, M_1 qualifies as intentional according to the definition. However, M_1 is clearly not an intentional sentence prefix. Thus, the definition does not provide a sufficient condition. The next example shows that the definition does not provide a necessary condition. Let M_2 be a syntactically primitive sentence prefix that is synonymous to the complex sentence prefix:

(______) or John believes that (______).

There are numerous sentences A such that $M_2(A)$ is not contingent. For example, since '$(2 + 2 = 4)$ or John believes that $(2 + 2 = 4)$' is necessary, $M_2(2 + 2 = 4)$ is necessary. Yet there is good reason for thinking that M_2 is an intentional prefix: '$(1 + 1 = 3)$ or John believes that $(1 + 1 = 3)$' is necessarily equivalent to 'John believes that $(1 + 1 = 3)$', which is undeniably intentional.[5] For another kind of example that shows that the definition does not provide a necessary condition, let M_3 be synonymous to the odd prefix:

9 believes that.

This prefix fails to satisfy the definition. E.g., $M_3(2 + 2 = 4)$ is not contingent since '9 believes that $2 + 2 = 4$' is not contingent. At the same time, M_3 would seem to qualify as an intentional prefix. To see why, notice that $M_3(A)$ expresses the proposition that 9 believes that A. Using only the law of existential generalization (EG), we can derive an undeniably intentional proposition, namely, the proposition that something believes that $2 + 2 = 4$. Consider a related example. Let M_4 be synonymous to the prefix:

This positron believes that.

Perhaps M_4 is a counterexample since, for all we know, 'This positron believes that $2 + 2 = 4$' is not contingent. For a final example, let M_5 be synonymous to the prefix:

John introspects that.

This is a counterexample since, e.g., 'John introspects that someone other than John is in pain' is not contingent.[6] However, M_5 is undeniably intentional.

What do these counterexamples show us about intentional relations? M_1 shows that there are certain non-intentional Cambridge relations whose logical behavior resembles that of genuine intentional connections. In a similar vein, M_2 shows that there are certain intentional Cambridge relations whose logical behavior fails to resemble that of genuine intentional connections. This suggests that the analysis of intentional relations should proceed in two steps. The first step is to define the concept of an intentional connection. And the second step is to define the concept of an intentional Cambridge relation in terms of this concept of intentional connection.[7] Once the first step is completed, the second step will be straightforward; for a Cambridge relation is intentional if its definition involves an intentional connection in a logically essential way. Thus, the crucial step is to define what it takes for a connection to be intentional.[8] But what do the remaining three counterexamples M_3, M_4, and M_5 show us? The lesson of M_3 is that universals cannot be intentionally connected to complex ideas; only individual particulars can. Next, M_4 indicates that not just any individual particular can be intentionally connected to complex ideas; only certain ones can. And finally, M_5 shows us that an individual particular cannot be intentionally connected to just any complex idea; there are epistemic limits. Nevertheless, for each intentional connection it is possible that at least some individual particulars are connected by it to at least some complex ideas.

Assembling the insights isolated over the last few pages, I offer the following analysis:

> A connection is intentional if and only if it can contingently connect an individual to a complex idea independently of the veracity of the idea.[9]

(An idea has veracity *iff*$_{\mathrm{df}}$ it is a true thought or a concept that applies to something.) The claim here is that all and only intentional connections have this special logical character. We intentional beings are distinctive in that we can stand in contingent connections to complex ideas independently of the veracity of the ideas. There are complex ideas with regard to which we can (but need not) believe, contemplate, decide, doubt, remember, want, etc., and we can (but need not) do this quite independently of whether these ideas are true or whether they apply to anything. In thought we can (but need not) do all sorts of things that are about, or at least purport to be about, objects in the world, even though these things we do need not in any relevant way correspond to the actual conditions of objects in the world. Thus, to think is to engage the dual possibilities of truth and falsehood, possibilities born for us through the weaving together of forms.

There are a number of candidate counterexamples to this analysis of the concept of intentional connection. Though many appear promising at first, none hits its mark. Still, some of them are of philosophical interest in their own right. For that reason, as well as for the reason of imparting a better feel for the analysis, I will explain why the best of these counterexamples fail.

(1) Take first the *predication relation*. True, this relation is a genuine connection, and it can connect individuals to complex ideas. However, it can connect an individual to a complex idea only when that idea has instances (e.g., the individual itself). Since the

analysis requires that intentional connections hold independently of whether the idea has instances, the predication relation is not a counterexample.[10] The *predication operation* fails to be a counterexample for a related reason. Though this operation is a genuine connection, when it holds it holds necessarily. But on the analysis it must hold contingently in at least some possible cases

(2) Certain *causally grounded dispositional relations* constitute another kind of potential counterexample. (Dispositions that are not causally grounded are not connections and so are not a source of counterexamples. For more on dispositional counterexamples, see note 24.) Consider the dispositional relation holding between particulars x and conditions y such that x is a hunk of salt having the disposition to dissolve whenever condition y obtains. This relation holds, for example, between the hunk of salt v in my hand and the condition c that v is submerged in the water in my glass. This relation is certainly not intentional. Yet it can hold contingently; it might not have held between v and c, for the water in my glass might have already been saturated with salt. And it holds between v and c whether or not c ever obtains, i.e., whether or not v is ever actually submerged in my water. (In this, its logical behavior resembles that of a subjunctive conditional.) Still, this relation is no genuine counterexample since it fails to meet the analysis on two counts. First, it is not a genuine connection; it is not one of the relations that fix the primary logical, causal, or phenomenal order of the world. To be sure, genuine causal connections do underlie this disposition, but these are connections that hold between particular salt molecules and particular water molecules (or between events involving particular salt molecules and events involving particular water molecules or between conditions involving particular salt molecules and conditions involving particular water molecules). And it is only in virtue of such prior causal connections that *ad hoc* dispositional relations between hunks of salt and conditions hold at all. The second count against this possible counterexample is that the relation holds only between particulars and conditions. But the analysis requires that the relation be able to hold between a particular and a *complex* idea (such as a thought). This reveals a weakness in all causally oriented counterexamples: qualities, connections, and conditions—the determinants of the causal order—are *simple* ideas, but it is *complex* ideas that typically are in the range of intentional relations.[11]

(3) The relation of *speaker meaning*—i.e., the relation holding between speaker x and complex idea y such that x means y by uttering something—might well be a connection. If it is, then all the requirements of the analysis of intentional connection are met. This is no counterexample, though, since the relation of speaker meaning is an intentional relation.

(4) Consider, finally, the relation of *utterance-token meaning*—i.e., the relation holding between utterance tokens and what they mean. This relation can hold contingently between a particular (namely, an utterance token) and a complex idea (namely, the meaning of the utterance token) independently of the veracity of the idea. Is utterance-token meaning truly a connection? It hardly seems so. An utterance token and the relevant complex idea are not related to one another just on their own; the intervention of a third element is required, namely, the intentional activity of thinking creatures. Not unless these creatures make utterances with appropriate intentions and beliefs do utterance tokens become related to the relevant complex ideas. Intending and believing are the genuine intentional connections; the relation between the utterance token and the complex idea that comes to be its meaning is entirely derivative. Unlike intending and believing, it plays no role in the primary causal and phenomenal order of the world.[12]

I have been unable to find better candidate counterexamples than these, and I have thus grown to be convinced that the analysis is free of all serious counterexamples. The thesis to which I wish to be committed, then, is that every candidate counterexample can be disqualified; at most minor alterations in the analysis might be called for. (E.g., minor adjustments might be called for in order to deal with someone's special doctrine about modality, existence, and time.[13]) Rather than attempting to tie up every loose end, I will stop here in hopes that the reader will be able to see how these adjustments would be made.

Perhaps what is most distinctive about this analysis of intentionality is that it is given entirely in terms of logic, specifically, the logic for qualities and concepts. Now someone might worry that the analysis is at bottom circular, for the theory of concepts is none other than the theory whose purpose it is to treat the logic for intentional matters. But this worry would be unfounded. Logic is logic regardless of its field of application. And the theory of qualities and concepts definitely counts as logic. Its primitive constants intuitively qualify as logical constants; they certainly are not smuggled in from psychology. (See also note 5 page 251.)

Having completed my analysis of intentionality, I am finally in a position to discuss Brentano's thesis of intentionality, which concerns the nature of the mental in general. Brentano's thesis will be the underlying theme of the next section.

49. Experience and the Mental

According to Brentano's thesis of intentionality, all and only mental phenomena are intentional, and moreover, no purely physical phenomenon is intentional. In a while I will take up the second half of this thesis; for the present my concern will be with the first half. Is it really true that intentionality is the mark of the mental? The counterexample that springs to mind is that of pure, uninterpreted experience—pure sensation or pure inner (emotional) feeling—as posited by traditional empiricists. Any such experience would certainly be a mental phenomenon, but it would not be *about* anything. Brentano of course wants to deny that there is any such thing as pure experience. However, Brentano puts forward the first half of his thesis as analytic. Therefore, this half of the thesis would be undermined if pure experience were merely possible for some beings or other, not necessarily human beings. In the face of this threat, we would be wise to have an analysis of the mental that is neutral with respect to the possibility of pure experience.[14] Since this is the goal, we must first get clearer about the sort of thing pure experience is supposed to be. I will begin by speculating about sensation and later on will take up inner feeling. (Throughout I will use 'sensation' to mean pure sensation and 'inner feeling' to mean pure inner feeling.) I should stress, though, that if pure experience is not possible, then all experience is intentional and no analysis of the mental is required beyond the previous analysis of the intentional.

Just as in the case of judgement, so in the case of sensation there are both relational and non-relational theories. Relational theories of sensation most often assert that sensation consists in standing in the relation *sensing* to special mental particulars, e.g., appearances, sense impressions, phantasms, sense data, sensa, etc. So, for example, if I were to gaze at a bright red tomato in normal well-lit conditions, then according to this theory I would come to stand in the sensing relation to a mental particular that is itself a red object. And similarly, if I were to hallucinate a bright red tomato, I would stand in the same sensing relation to the same sort of mental red particular.

Realists find it incredible that there are any such special mental particulars. They find it incredible, e.g., that, when I hallucinate something red, there is a real object sensed by me that is actually colored red. This reaction of the realists is certainly reasonable. Yet many realists carry it to an overreaction. They go on to conclude that sensation cannot in any way consist in standing in relations to objects and, therefore, that the non-relational theory of sensation must be right. The adverbial theory of sensation is one instance of a non-relational theory that is arrived at by this route. Versions of the adverbial theory have been advocated by Ducasse, Chisholm, Ayer, and, at times, Russell. On Chisholm's version of the adverbial theory, for example, my experience in the above two cases should be described by means of an explicitly adverbial construction, e.g., 'I am appeared to *redly*' or 'I sense *redly*'. All suggestion of the relational theory of sensation is thereby avoided.

Just as there is an analogy between non-relational theories of sensation and non-relational theories of judgement, there is also an analogy between adverbial theories of the language for sense experience and operator (and prefix) theories of the language for intentionality. In the last section I showed that the matter of generality created logical difficulties for non-relational theories of judgement and operator (and prefix) theories of the language for intentionality. It is predictable, therefore, that non-relational theories of sensation and adverbial theories of the language for sense experience should also run into logical difficulties on the issue of generality. For example, on the adverbial theory how are we to express the thought that the sense experience of one creature u is exactly like that of another creature v? Perhaps one could express this thought by means of some baroque higher-order adverbial theory. E.g., one might attempt to treat adverbs as a special new category of singular terms whose semantical correlates are "ways of experiencing". Accordingly, one might attempt to represent the above thought about creatures u and v in something like the following manner:

$$(\forall F)(F(\mathrm{Senses}(u)) \equiv F(\mathrm{Senses}(v))).$$

But notice how close this comes to being just a higher-order version of the relational theory of sensation, the theory that the adverbial theory is designed to avoid. How much more direct it is to express the thought about creatures u and v by means of an explicitly relational theory:

$$(\forall x)(u \text{ senses } x \equiv v \text{ senses } x)$$

i.e., u senses whatever v senses, and conversely.[15] For this and related reasons I am inclined to conclude that, if there is such a thing as pure sensation, then the best theory of it is the relational theory.

Once one adopts the relational theory, however, one is obliged to identify the sort of objects that are in the range of the sensing relation. What are they? In order to help find the answer to this question, let us consider again the thought that the sense experience of creature u is exactly like that of creature v. Intuitively, this thought could be true. So for the purpose of discussion, let us suppose that it is. Then, given the above relational analysis of the thought, it follows that, in this example at least, the objects of the sensing relation are not private objects.[16] Next, let us suppose that creature u or creature v or both u and v are hallucinating some or all of the time. Even in this case, it still seems that the thought that the sense experience of u is exactly like that of v could be true. So for the purpose of discussion, let us again suppose that it is. Then, given the above relational analysis of the thought, it follows that at least in this example the objects of the sensing relation are not ordinary physical particulars. Therefore, if in this example the objects of the sensing relation were particulars at all, they would have to be some kind of public particular that has no actual location, no actual causal efficacy, etc. This, however, sounds rather like a kind of particular that simply could not exist. It offends virtually every realistic intuition we have.

Barring such unacceptable particulars, one has no alternative but to conclude that at least in this example some objects of the sensing relation are universals, namely, certain appropriate qualities and conditions.[17] Such qualities and conditions are called *sensible qualities* and *sensible conditions*. By allowing at least some of the objects of the sensing relation to be sensible qualities and sensible conditions, one is still able, as desired, to hold a version of the relational theory of sensation. This version of the theory permits the objects of sensation to be public objects, as desired. For qualities and conditions are public objects. And at the same time, this version of the relational theory avoids the realist objections that are so damaging to the traditional sense-data theory. According to the present version of the relational theory, one can, for example, sense the color red quite independently of whether any particular is actually colored red (as in hallucinations or dreams).[18] Likewise, given that one can sense conditions, one can do so quite independently of whether they actually obtain. For example, given that one can sense the condition that something colored red is surrounded by something colored blue, one can do so quite independently of whether there actually is something colored red surrounded by something colored blue. In general, on this version of the relational theory, the sensing relation is such that it can hold between an individual and a quality or condition independently of whether the quality or condition is *concretized*, i.e., independently of whether it either has an instance or obtains.

Now whenever the sensing relation holds in the way just indicated between some individual and some quality or condition, it of course does so only contingently. In addition, if there is such a thing as pure sensation, it would seem that the sensing relation is a genuine connection, rather than a mere Cambridge relation. But if so, look how close this comes to the analysis of intentional connections. An intentional connection is one that can contingently connect an individual to a complex concept or thought independently of the veracity of the concept or thought. On analogy, the sensing relation is a connection that can contingently connect an individual to a quality or condition independently of whether the quality or condition is concretized. Just as in thinking one engages the possibility of falsehood, so in sensing one engages the possibility of illusion.

With these remarks about sensing in mind, let us now turn to the topic of inner feeling, the kind of feeling traditionally thought to be associated with emotion. If the bodily-sensation theory of inner feeling were correct, then our job would be done, for inner feeling in that case would be just a species of sensing. If, on the other hand, this theory is not correct, then feeling must be taken up separately. From a logical point of view feeling does seem unlike sensing; whereas an individual can sense a sensible quality independently of whether anything has the quality, an individual can feel, say, a pure emotional quality only if he himself has the quality. For example, whereas one can sense red independently of whether anything is red, one can feel sad only if something—oneself—is sad at least momentarily.[19] This might suggest that the relations of feeling and sensing are unrelated. Perhaps this is so. However, there is an attractive alternative which I will suggest.[20]

According to this alternative, sensing and feeling are not distinct basic modes of pure experience. There is in fact only one basic mode of pure experience, namely, *pure experiencing* itself. Sensing and feeling differ only in their objects. Sensing is the relation that results from restricting the range of the experiencing relation to sensible qualities and sensible conditions, i.e., qualities and conditions that can be experienced independently of whether they are concretized. And feeling is the relation that results from restricting the range of the experiencing relation to reflective qualities, i.e., qualities that an individual can experience only if they are qualities of that individual at least momentarily.[21]

On the view of the mind that is emerging there are two basic types of mental phenomena: pure experience and thinking (where the latter is taken to include all that is intentional).[22] The difference between them is that in pure experience we are typically connected to qualities and conditions and in thinking we are typically connected to complex concepts and thoughts. From a purely logical point of view, the difference between qualities, connections, and conditions, on the one hand, and thoughts and complex concepts, on the other, comes down to one of logical form. The former are simple with regards to logical form; the latter are complex. However, all of these entities, whether simple or complex, are intensional entities, and indeed they are the only intensional entities. That is, all and only these entities are ideas, as intensional entities are called in the theory of qualities and concepts. This suggests a unified analysis of what a mental connection is:

> A connection is mental if and only if it is—or is necessarily included in—a connection that can contingently connect an individual to an idea independently of whether that idea is realized.[23]

(An idea is realized *iff*$_{\mathrm{df}}$ it is a true thought, a concept that applies to something, a quality that has instances, a connection that connects something, or a condition that obtains.)

Intuitively, what is distinctive about a mental being is this. He can stand in contingent connections to ideas, and he can do so in such a way that it is not crucial how these ideas actually correspond to, or show up in, the world unless these ideas reflect special aspects of his own mental conditions, in which case they are realized in him. In so connecting an individual to ideas, mental connections thus provide him with a highly adaptive kind of "window on the world" whose reliability typically is variable, one exception being when it reflects certain special aspects of the individual's own mental conditions. Although non-mental individuals are connected to ideas in various ways, they are never connected to ideas in these unique ways. And this makes all the difference.

The candidate counterexamples that come to mind are all variants of those facing the analysis of intentionality, and they can be disqualified for corresponding reasons. One kind of candidate counterexample deserves special comment, however: namely, causally grounded dispositional relations. Consider the relation holding between x and y such that x is disposed to be activated by y—i.e., the relation holding between x and y such that y (dispositionally) activates x. Let us suppose what might well be false, that this relation is a connection. A critic of the analysis might claim that this relation can contingently connect a particular to a quality independently of whether the quality has any instances. He might claim, for example, that it can contingently connect a particular photoelectric cell to the color red independently of whether anything is actually colored red. If the critic is right, then the relation would be a true counterexample. I would dispute what the critic claims, however. In order for one thing to *dispositionally* activate another, it must be the kind of thing that can *actually* activate that other thing. But what actually activates a photoelectric cell, for example, is not a color itself but rather particular electromagnetic waves, which might be instances of the color. So the color red is not the kind of thing that can dispositionally activate a photoelectric cell, contrary to what the critic claims. True, a photoelectric cell can bear derivative Cambridge relations to a quality by being connected to instances of the quality. But there is no way for a photoelectric cell and a quality to stand in a genuine connection of the sort the critic imagines; the mediation of an instance of the quality is required. What is special about pure experience (if it truly exists) is that in it we are connected to sensible qualities without the need for the mediation of instances of the quality; we can sense red in hallucinations, illusions, and dreams even if nothing is actually colored red.[24]

This analysis of the concept of a mental connection is stated entirely within the logic for qualities and concepts. Nothing but fundamental logical relations are appealed to: the predication relation, the thought-building operations (conjunction, negation, existential generalization, etc.), and the associated condition-building operations. Furthermore, the analysis, unlike the one envisaged by Brentano, does not rule out the possibility of pure experience. Thus the analysis holds independently of the first half of Brentano's thesis of intentionality, i.e., independently of the conjecture that all and only mental phenomena are intentional. This is fortunate, for there is a barrier blocking a full defense of the first half of Brentano's thesis: how could one ever demonstrate that pure experience is not at least possible for some being or other?

The second half of Brentano's thesis—i.e., the conjecture that no purely physical phenomenon is intentional—fares better than the first half, for there is no comparable barrier to its defense. Let us now look at this half of his thesis in relation to the issue of materialism. In order to state the doctrine of materialism, one must pay attention to the question of what it takes for an object to be physical in the materialists' sense.

Suppose that a particular is connected by some connection to some object. Then for brevity I will say that the particular *has* the connection. For example, every particular stands in the predication relation to the qualities that qualify it and to the concepts that apply to it. (In these cases the predication relation holds sometimes necessarily, sometimes contingently.) Thus, since the predication relation is a connection, every particular has the predication relation as one of its connections. There are also certain other logical connections that every particular has. Some of these the particular has necessarily. The following example involves one such necessary logical connection. Consider the particular x and an arbitrary property $|Fy|_y$ and the associated condition $|Fx|^x$. The condition-building operation of predication connects the three objects $|Fy|_y$, x, $|Fx|^x$, and it does so necessarily. Thus, the particular x has this necessary logical connection as one of its connections. In much the same vein, a particular also has necessary logical qualities, i.e., logical qualities that necessarily qualify the particular. Thus, every particular has various necessary logical qualities and connections;

in addition, every particular also has the predication relation as a purely logical connection.

Now materialism requires that, besides the purely logical qualities and connections of the sorts just described, the only qualities and connections that any particular has are physical qualities and connections. Particulars that are like this may be called *purely physical.* And those phenomena involving qualities and connections each of which is physical likewise may be called *purely physical.*[25] Materialism then is the doctrine that all particulars are purely physical and, in turn, all phenomena are purely physical.

Materialists historically have been unclear about the meanings of their basic terms; in particular, a fog surrounds the key term 'physical'. I believe that the only way one can remedy this situation is to analyse the basic concepts of materialism (and physicalism) within a purely logical theory; i.e., one must employ a strategy akin to the one I employed in analysing intentionality and the mental. Such an analysis can make use of the following necessary condition. Genuine physical connections can hold only between particulars and particulars, and perhaps between particulars and locations, particulars and times, particulars and stuffs, locations and locations, times and times, stuffs and stuffs, etc.;[26] they cannot hold between particulars and complex ideas. If a genuine connection holds between particulars and complex ideas, then it is not a physical connection. Whoever maintains otherwise would appear to have forgotten a category difference between physical and mental connections.

According to my analysis of intentionality, an intentional connection is one that can contingently connect a particular to a complex idea independently of the idea's veracity. Two conclusions follow. First, no intentional connection is a physical connection. Secondly, no intentional connection is the sort of special logical connection (characterized above) that the materialist permits purely physical particulars to have. From these two conclusions a third follows. For any intentional connection x, any particular y, and any complex idea z, the intentional phenomenon that x connects y to z is not a purely physical phenomenon. And this conclusion can be generalized, yielding the conclusion that no intentional phenomenon is a purely physical phenomenon. This is none other than the second half of Brentano's thesis of intentionality.

The same argument also works for mental connections and mental phenomena generally. On my analysis, mental connections are necessarily included in connections that can contingently connect a particular to an idea (i.e., to an intensional entity) independently of that idea's being realized. No physical connection is like this. Thus, no mental connection is a physical connection. At the same time, no mental connection is the sort of special logical connection that the materialist permits purely physical particulars to have. It follows that, for any mental connection x, any particular y, and any idea z, the mental phenomenon that x connects y to z is not a purely physical phenomenon. Generalizing, one may conclude that no mental phenomenon is a purely physical phenomenon. This last conclusion is the main consequence that Brentano wanted to derive from his full thesis of intentionality. Fortunately, it is obtained here without appeal to the first half of Brentano's thesis.

From the conclusions reached in the preceding paragraph it follows that, if there are beings who in fact have mental connections to things, then there are beings who are not purely physical and there are phenomena that are not purely physical. Now, *à la* Descartes, I know directly that I am thinking thoughts. I also know directly that I am sensing smells, sounds, etc. and that I am feeling emotions. However, I can be thinking thoughts and having sensations and feelings only if I am standing in mental connections to things. And since I have mental connections to things, I am not purely physical, and phenomena involving my mental connections are not purely physical. Hence, if I am identical to my body, then my body is not purely physical, and phenomena involving my body's mental connections are not purely physical. And if I am not identical to my body, then phenomena involving nothing but my body are not identical to phenomena involving me, phenomena such as my thinking, my sensing, and my feeling. Either way, I am not identical to a purely physical object, and phenomena involving my mental connections are not identical to purely physical phenomena. Finally, all the foregoing conclusions about mental phenomena hold *mutatis mutandis* for mental events. Whether or not I am identical to my body, mental events involving me are not purely physical.

Let us consider now the Chisholm/Sellars dispute on intentionality and the mental which I described briefly in the previous section. In this dispute Chisholm, echoing Brentano, maintained that no intentional sentence could be analysed without making further appeal to intentional vocabulary, and Sellars, espousing his own special version of materialism, maintained that every intentional sentence could be analysed without appealing to any intentional vocabulary. I will argue that, given my analysis of the concept of an intentional connection, both Chisholm and Sellars were in error. Consider Chisholm's position first. Given my analysis, there are at least some intentional sentences, e.g., 'Someone stands in an intentional connection to some thought', that can be analysed without intentional vocabulary; purely logical vocabulary suffices here. So, given my analysis, Chisholm's position is too strong. In fact, Chisholm's position that no intentional sentence can be analysed without appeal to further intentional vocabulary is nearly inconsistent with his view that the concept of an intentional sentence has a purely logical analysis. Next consider Sellars' doctrine that all intentional sentences can be analysed without appeal to intentional vocabulary. To see what the problem is with Sellars' doctrine, consider an intentional connection that connects a given individual (e.g., me) to a thought (e.g., the thought that I think). Let the 2-place predicate 'T' express this intentional connection; let the name 'a' name the individual, and let the name 'b' name the thought. Now perhaps materialists could *describe* in non-intentional vocabulary the truth expressed by the sentence '$T(a, b)$'. Nevertheless, they could never produce a non-intentional sentence that *expresses* this truth. The reason goes as follows. Sentences, I have shown, express thoughts. However, thoughts conform to conception 2. Therefore, by the laws of the logic for conception 2, it follows that the only sentences that could ever express the same thing as the sentence '$T(a, b)$' are sentences having the logical form '$F(c, d)$' or '$G(c)$', where 'F' and 'G' are primitive predicates. But in this case 'F' would have to express just the intentional relation expressed by the intentional predicate 'T', and 'G' would have to express just the complex intentional concept expressed by the intentional open sentence '$T(x, b)$'. So 'F' and 'G' would themselves have to be intentional predicates. And so '$F(c, d)$' and '$G(c)$' would have to be intentional sentences. Thus, it is impossible to express in non-intentional vocabulary the truth expressed by the original intentional sentence. It follows, therefore, that the materialist inevitably leaves something out.

I will close my discussion of the mental with a brief comment on minds and machines. Suppose that we should one day design and build a machine that performs physically as we do, both behaviorally and mechanically. A natural question to ask is whether the machine has a mind. According to behaviorism and materialistic versions of functionalism, this is just the question of whether the machine behaves or functions physically as we do. But *ex hypothesi* we already know that it does; *that* is not our question. We want to know something else, namely, whether the machine actually functions mentally. But this is to say, we want to know whether it stands in genuine mental connections to things. For intuitively, a thing functions mentally if and only if it stands in mental connections to things. It is not enough that it should behave or function physically as if it were mentally connected to things. Research on minds and machines that disregards this difference is likely to reach misleading conclusions about the basic nature of the mind.[27]

A closely related question—and one with the greatest moral significance—is whether a machine that behaves and functions physically as we do is conscious, i.e., whether it is aware of anything. Before one tries to settle this question, one should try to get clearer about what consciousness is; i.e., one should try to say what it is for a being to be aware of something. I will attempt to do this in my final, rather speculative section.

50. Consciousness

Suppose for a moment that all mental connections are *conscious connections*; i.e., suppose that, necessarily, whenever an individual is mentally connected to something, he is also conscious of that thing. In this case, since I have already given a purely logical definition of the concept of a mental connection, it would be a straightforward affair to obtain a purely logical definition of consciousness also. Perhaps this is all there is to it. However, if certain commonly held psychological theories are correct, there are mental connections that are not conscious. (Examples might be standing belief, standing desire, unconscious desire, unconscious decision, etc. Let us call such mental connections *non-conscious*.) If there truly are non-conscious mental connections, then some other strategy for defining consciousness is in order. Fortunately, one can sidestep the complicated theoretical issue of whether there are non-conscious mental connections, for there is another way to define consciousness that is neutral with regard to how that issue is settled. The key to the definition is the *unity of consciousness*.

At a given moment I might be sensing one thing, feeling another, thinking a third, and desiring and deciding still others. Yet despite the fact that several specialized conscious connections are all operating at once, my consciousness is not fragmented into a corresponding variety of exclusive programs on competing channels. I have a unified awareness of all the objects of these specialized conscious connections. What accounts for the unity?

The answer is that the consciousness relation—the relation being conscious of or being aware of—is itself a conscious connection whose unique global operation produces this unity. Thus, if an individual has a pure experience of something, then he will also stand in the consciousness relation to that thing; if an individual believes, wants, or decides something consciously, then again he will stand in the consciousness relation to it, and so on for any specialized conscious connection.[28] The consciousness relation is that mental connection whose operation must have this global character. The problem is to get at this character of the consciousness relation without circularity, i.e., without explicitly alluding (as I just did) to the distinction between conscious and non-conscious mental connections.

To be sure, if every mental connection were a conscious connection, the problem would be easy to solve: consciousness would simply be the maximal mental connection, i.e., the mental connection that is the union of all mental connections.[29] But I am looking for a definition that is compatible with psychological theories that entertain non-conscious mental connections, so another approach is needed.

Notice that the non-conscious mental connections have something in common. In each case, they mimic functionally the operation of a conscious counterpart, and moreover, in each case this conscious counterpart is one of the specialized conscious connections such as conscious belief, conscious want, conscious decision, etc. The consciousness relation itself is not the conscious counterpart of a non-conscious mental connection. What would that non-conscious connection be, unconscious consciousness, unaware awareness? The theoretical purpose served by non-conscious mental connections is to constitute a non-conscious functional analogue of conscious mental processes. But the consciousness relation has no function that could possibly show up when mental processes are looked at mechanically as if there were no consciousness involved. All it does is to produce a unified consciousness. There cannot be a non-conscious connection with that function since non-conscious mental connections are, as we say, non-conscious. (This is why the consciousness relation never appears in functional psychologies based on the information-processing model.)

The conscious connections with non-conscious counterparts are those that do have mechanically recognizable functions. Conscious belief, conscious want, conscious decision, etc. are like this.[30] There are, however, certain other specialized conscious connections—such as attending, concentrating, meditating, contemplating[31]—that are not like this. Akin to the consciousness relation, they lack any apparent function when mental processes are viewed mechanically as though no consciousness were involved. Non-conscious attending, non-conscious concentrating, non-conscious meditating, non-conscious contemplating—these are useless Cambridge relations if any are. The only immediate function of relations like attending and meditating is to alter the quality of consciousness. (Thus I will call relations of this kind *qualitative conscious relations.*) For example, if an individual attends to or concentrates on something, then he will thereby acquire a *keen* awareness of that thing; if he meditates on or contemplates something, then he will therein have a *heightened* consciousness of it.[32]

Consider now the conscious connections that do have non-conscious counterparts. Notice that these conscious connections have special category limitations on their ranges. For example, it is impossible to believe consciously a color, a taste, or a smell; to want consciously a number; to decide consciously oneself, etc.[33] Naturally, the special category limitations on these conscious connections are inherited by their non-conscious counterparts. So, for example, just as it is impossible to believe consciously a color, a taste, or a smell, it is also impossible to believe non-consciously a color, a taste, or a smell; and so on.

Let us recall, however, that the consciousness relation is a global mental connection in the sense that, if an individual stands in any conscious connection to an object, then he will also be conscious of that object. The consciousness relation thus lacks the special category limitations of the sort in force for the mechanically significant conscious connections and their non-conscious counterparts; in the sense that its range is free of these limitations the consciousness relation is *transcendental*.[34] I can be conscious of colors, sounds, smells, numbers, and I think that I am at the present conscious of myself and the present.[35] The qualitative conscious relations (attending, etc.) are like the consciousness relation in this. They too lack the special category limitations on their ranges, and in this sense they too are transcendental mental relations. So in general, if it is possible for an individual to stand in any conscious or non-conscious mental connection to an object, then it is also possible for an individual to be conscious of that object, and likewise, it is possible for some individual to attend to (concentrate on, contemplate, meditate on) that object. Now what is unique about the consciousness relation is that it is the maximal mental connection whose range has this transcendental character. That is, it is the mental connection that is the union of all transcendental mental connections, including itself. This is what is global about the consciousness relation.

Notice that this captures the unique global character of consciousness without circularity, i.e., without mention of the distinction between conscious and non-conscious mental connections. Thus, I am finally in a position to state my definition:

> consciousness $=_{\mathrm{df}}$ relating by means of transcendental mental connections.

That is,

$$|x \text{ is conscious of } y|_{xy} =_{\mathrm{df}}$$
$$|(\exists z)(z \text{ is a transcendental mental connection} \;\&\; \langle x, y\rangle \,\Delta\, z)|_{xy}.$$

Inasmuch as the consciousness relation is the mental connection having this unique global character, its operation is what produces the unity of consciousness.[36]

From the definition of consciousness, definitions of conscious mental quality and connection follow immediately. Now, each new category of qualities and connections generates the possibility of a new category of beings and, thus, the possibility of a new category of associated phenomena. So it is with the category of conscious mental qualities and connections; it too generates the possibility of a new category of beings and associated phenomena. Indeed, that possibility is actual, for we are among the new beings.

Notes

Introduction

1. Pp. 75 ff., Nelson Goodman, *Fact, Fiction, and Forecast.*
2. 'A Formulation of the Logic of Sense and Denotation' and 'Outline of a Revised Formulation of the Logic of Sense and Denotation'.
3. 'Intensional Isomorphism and Identity of Belief'. This conception of synonymy is assessed in §19 below; evidently it is an outgrowth of Church's effort to find a formally adequate resolution to the paradox of analysis.
4. Axioms 63–8, 'Outline', part two.
5. It is very important to realize that logical validity and epistemic justification also fall under conception 2.
6. See §14 for an illustration.
7. What I call Russell's theory is a synthesis of positions Russell took in the writings of his early period.
8. The generous grades on desiderata 8 and 9 are meant to imply only that the various resolutions of the paradoxes are successful in avoiding the paradoxes; it is doubtful that an ideal resolution has yet been found.
9. This argument is adapted from George Myro's important paper 'Aspects of Acceptability'.

Chapter 1

1. Whether this quantifier logic should be first-order or higher-order is not relevant at the moment. That question will be taken up in §10 and again in chapter 4.

   It will be noticed that in a higher-order language having sentential (i.e., propositional) variables (I) can be represented by treating 'is necessary', 'is true', and '*x* believes' as operators that take sentences into sentences. However, this is so only because in such higher-order settings there is no hard distinction between predicates and operators. Thus, the conclusion in the text stands. By the way, in chapter 4 I propose to treat sentences such as '*a* is red' as having the form '*a* is *b*' where the copula 'is' is a 2-place logical predicate expressing the predication relation. If this treatment is right, perhaps '*a* is necessary' and '*a* is true' should be treated analogously. If so, then strictly speaking 'is necessary' and 'is true' would not be predicates. Still, this does not affect the substance of my claim that 'is necessary' and 'is true' are predicates, for like ordinary predicates, 'is necessary' and 'is true' would still combine with a singular term to yield a sentence (open or closed). This is the only point that is needed for the succeeding steps in my argument in the text.
2. I take up the question of the definability of the bracket notation in §37.
3. Incidentally, the conclusion that 'that'-clauses should be treated as defined or undefined singular terms is compatible with all approaches considered on the chart in §4 except for the approaches of Carnap, Hintikka, and Davidson and the Quinean primitive-predicate approach.

4. The possibility of externally quantifiable occurrences of variables is not allowed in Quine's original bracket notation.
5. For example, it is intuitively valid that if it is true that *A*, then there is something that is true. It would be irrational to deny this and to hold instead that the antecedent could be true and the consequent false. This shows that 'is true' is a predicate satisfied by entities. Given this, the best theory of what makes 'It is true that *A*' true is that the constituent predicate 'is true' is satisfied by an appropriate entity. One could hold otherwise only by disunifying his treatment of truth, and what good reason could there be for that? Given the fact that what makes 'It is true that *A*' true is that the predicate 'is true' is satisfied by an appropriate entity, it would only be perverse to deny that the entity is other than one semantically correlated with the singular term 'that *A*'.

   Of course, one wants a semantical account, not just of 'that'-clause sentences concerning truth, but of an open-ended list of 'that'-clause sentences, sentences concerning validity, provability, evidence and epistemic justification, explanation, all the various psychological attitudes, meaning, assertion, the modalities, causation, probability, counterfactuality, moral prescription, etc. In view of the open-ended character of this list and in view of the fact that all forms of 'that'-clause sentences may be embedded in one another arbitrarily many finite number of times, a unified, general account demands the apparatus of quantification and cross reference in connection with 'that'-clauses.
6. The sentence (a) 'There is a language *S′* such that Seneca wrote as a sentence of *S′* words whose translation from *S′* into English is "Man is a rational animal"' is a typical nominalistic analysis of sentence (1) 'Seneca said that man is a rational animal', which contains a 'that'-clause. Church criticizes this analysis,

   > For it is not even possible to infer (1) as a consequence of [(a)], on logical grounds alone—but only by making use of the item of factual information, not contained in [(a)], that 'Man is a rational animal' means in English that man is a rational animal.
   >
   > Following a suggestion of Langford we may bring out more sharply the inadequacy of [(a)] as an analysis of (1) by translating into another language, say German, and observing that the two translated statements would obviously convey different meanings to a German (whom we may suppose to have no knowledge of English). The German translation of (1) is (1′) *Seneca hat gesagt, dass der Mensch ein vernünftiges Tier sei*. In translating [(a)], of course 'English' must be translated as 'Englisch' (not as 'Deutsch') and '"Man is a rational animal"' must be translated as '"Man is a rational animal"' (not as '"*Der Mensch ist ein vernünftiges Tier*"').

   (See Church, 'On Carnap's Analysis of Assertion and Belief'.) Incidentally, let us suppose with Quine that there are certain epistemological difficulties in determining which one in a class of candidate translations is correct. Still, *epistemological* difficulties in making a distinction do not in general entitle one to draw the *ontological* conclusion that the distinction does not exist. Quineans have been very hard put indeed to show why this generalization does not apply to the distinction between correct and incorrect translations.
7. For example, nominalistic approaches appear to provide no adequate treatment of prelinguistic intentional states such as those of infants and intelligent lower animals. Relatedly, they appear to provide no adequate treatment of complex intentional states of people who know no single language but only know fragments of several distinct languages—e.g., people who philosophize in Greek and make love in French. Carnapian and Quinean syntactical analyses are, in

addition, faced with a special difficulty of their own. They seem to require new *ad hoc* expressions such as 'believes-true A as a sentence of language L'. But what do such new expressions mean? It would seem that this question can be answered only if appeal is made to the theory of propositions.

The language-of-thought theory is a nominalistic account which avoids many of the above difficulties. But this theory is caught in the following dilemma. Either it must be construed as a form of representationalism (in the sense of §42); in this case the relation between thoughts and what they correspond to in the world remains mysterious. Or it must take the radical position that no one's beliefs ever have any real content; in this case the theory would seem to be logically self-defeating (since *it* would presumably be believed by its proponents). Either way, the theory has a deep problem not found in the theory of PRPs. Of course, a proponent of the theory might try to avoid this verdict by defining the relation holding between thoughts and what they correspond to in the world. However, these efforts all end in failure unless intensions are re-introduced: behavioristic definitions cannot be sufficiently discriminating, as Quine's indeterminacy argument shows in effect; physiologically oriented definitions must fail, for among other reasons, because of the open-endedness of the possible physiological bases of thought; and functional definitions run into the difficulties mentioned in note 27 of chapter 10, difficulties which can be surmounted only by appealing to intensions of the sort posited in the theory of qualities and concepts.

8. I emphasize again that this conclusion takes no stand on the question of whether this intensional abstraction operation is defined or undefined. Both positions will be considered later in the book.
9. Leon Henkin has proved a quasi-completeness result for higher-order quantifier logic: when the language of higher-order logic is interpreted with what Henkin calls general models (as opposed to the usual standard models), higher-order logic is complete. (Every standard model is a general model, but not conversely.) Advocates of the higher-order approach sometimes point to Henkin-style quasi-completeness results to try to show that on the issue of completeness the higher-order approach is not inferior to the first-order approach. However, it is a matter of considerable controversy whether these results warrant any such philosophical conclusion.
10. Transcendental predicates produce comparable difficulties for many of the syntactic theories that are categorial in style.
11. It should be recalled that in Quine's notation, unlike the notation I have proposed, externally quantifiable variables are not permitted to occur within bracketed expressions $[A]_\alpha$. (Strictly speaking, Quine uses $\alpha[A]$ where I use $[A]_\alpha$, but this notational variant is of no significance.) Also, even though Quine would be willing provisionally to interpret $[A(v_1, \ldots, v_i)]_{v_1 \ldots v_i}$ as a term that denotes an i-ary intensional entity, he in the end would want to interpret it as a term that denotes the formula $A(v_1, \ldots, v_i)$ itself. This nominalistic interpretation, however, need not concern us here. The issue I want to focus on is syntactic, not semantic.
12. I do not approve of this treatment. For, given the arguments of chapter 5, sets do not really exist, and sequences turn out to be just a special kind of *de re* property or *de re* relation-in-intension. (*De re* PRPs are those that we would naturally denote with intensional abstracts containing externally quantifiable variables.) The modified Quinean approach to the logic of *de re* PRPs is thus caught in a vicious regress, one fraught with internal technical inconsistencies.
13. This limitation is another count against the modified Quinean treatment itself.
14. In 'Intensional Logic in Extensional Language' Charles Parsons constructs a Fregean system which can represent quantifying-in (and which can satisfy

Davidson's finite-learnability requirement). However, Parsons accomplishes this by adjoining to the system an *ad hoc* device whose semantic force is to associate with each entity a special "rigid" concept, i.e., a special essential individuating concept of that entity. But such concepts, if they are credible at all, are just singular concepts, i.e., the sort of concepts that arise from singular predications. (E.g., the special individuating concept of me is surely just the concept of being identical to me, i.e., $[x = y]_x^y$ where $y =$ me.) Ideally, however, a logical theory should treat singular concepts and singular predication directly; only then will one be able to lay bare the logic for the special "rigid" concepts posited by Parsons. (This is what is done by means of my bracket notation and the semantics for it (see §13).) Singular concepts and singular predication, moreover, are not even countenanced by Frege's philosophy (though they are by Russell's). So once again quantifying-in would seem to be representable (and Davidson's learnability requirement would seem to be satisfiable) only by retreating from Frege's original view (and by taking up a neo-Russellian position instead).

Chapter 2

1. Nothing prevents us from adjoining primitive functional constants to L_ω, but that would require enriching the algebraic model structures (see §14) by adding operations for application of function to argument and relativized applications of function to argument.
2. These operations will be precisely defined in the next section. However, it might be helpful to describe provisionally the relativized predication operations, which are more difficult to understand than the others. An intensional abstract $[A]_\alpha$ binds those free variables in the embedded formula A that occur among the variables α. What is special about relativized predications is that some of the variables bound by the intensional abstract $[A]_\alpha$ occur free in an intensional abstract occurring within the embedded formula A. So, for example, the intensional abstract $[F[Gy]^y]_y$ binds the variable y that occurs free in the intensional abstract $[Gy]^y$ occurring within the embedded formula $F[Gy]^y$. And more generally, the abstract $[F[A]_{v_1 \ldots v_m}^{u_1 \ldots u_k}]_{u_1 \ldots u_k}$ binds the variables $u_1, \ldots, u_k$ in the embedded formula $F[A]_{v_1 \ldots v_m}^{u_1 \ldots u_k}$. This abstract is the k-ary relativized predication of $[Fx]_x$ of $[A]_{v_1 \ldots v_m u_1 \ldots u_k}$.
   There is an alternate strategy for dealing with relativized predications. Instead of denumerably many relativized predication operations having two arguments, one posits a single predication operation having three arguments, the additional argument serving to code the number of variables to be relativized in a particular application of the operation. This alternate strategy is sketched in my 'Theories of Properties, Relations, and Propositions'.
3. For more on the algebraic approach to extensional logic (without abstraction operations), see Henkin, Monk, and Tarski, *Cylindric Algebras*. Incidentally, the first seven operations also have a close relationship to the syntactic operations isolated in Quine's 'Variables Explained Away'.
4. Strictly speaking $\mathscr{P}$ is a prelinear order on $\mathscr{D}$.
5. Readers inclined to view $\mathscr{K}$ as a vestige of possible worlds should see p. 209 f. By the way, the truth values (T and F) may be defined in many ways; they may be identified respectively with $\mathscr{D}$ itself and with the null set, for example.

6. As things stand $\mathscr{D}$ is not closed under these operations. For example, Neg is not defined for elements of $\mathscr{D}_{-1}$. To close $\mathscr{D}$ under Neg, one could identify Neg(x), for $x \in \mathscr{D}_{-1}$, with some arbitrary element of $\mathscr{D}$. The same goes for the other operations. By the way, conservative Platonists might wish to modify clause (8.0) as follows: $\text{Pred}_0 : \mathscr{D}_i \times (\mathscr{D} \sim \mathscr{D}_{-1}) \to \mathscr{D}_{i-1}$. This modification rules out the possibility of particulars being genuine subjects of predications.
7. In general, for $i \geq 1$ and $j \geq k \geq 1$, $\text{Pred}_k : \mathscr{D}_i \times \mathscr{D}_j \to \mathscr{D}_{j+k-1}$. In my informal remarks in the previous section Pred_0 is what I called absolute predication; Pred_1, unary relativized predication: Pred_2, binary relativized predication; ...; Pred_k, k-ary relativized predication.
8. In general,

$$8.k \quad \langle x_1, \ldots, x_{i-1}, y_1, \ldots, y_k \rangle \in H(\text{Pred}_k(u, v)) \equiv \langle x_1, \ldots, x_{i-1}, \text{Pred}_0(\ldots \text{Pred}_0(\text{Pred}_0(v, y_k), y_{k-1}), \ldots, y_1) \rangle \in H(u)$$

where $u \in \mathscr{D}_i$, $i \geq 1$, and $v \in \mathscr{D}_j$, $j \geq k \geq 1$. The following will help to illustrate the behavior of the predication operations Pred_0, Pred_1, ...:

Since $\text{Pred}_0([Fx]_x, [Guv]_{uv}) = [F[Guv]_{uv}]$, clause (8.0) insures that $H([F[Guv]_{uv}]) = \text{T} \equiv [Guv]_{uv} \in H([Fx]_x)$.

Since $\text{Pred}_0([Fxy]_{xy}, [Guv]_{uv}) = [Fx[Guv]_{uv}]_x$, clause (8.0) insures that $x \in H([Fx[Guv]_{uv}]_x) \equiv \langle x, [Guv]_{uv} \rangle \in H([Fxy]_{xy})$.

Since $\text{Pred}_1([Fx]_x, [Guv]_{uv}) = [F[Guv]^v_u]_v$, clause (8.1) insures that $v \in H([F[Guv]^v_u]_v) \equiv [Guv]^v_u \in H([Fx]_x)$.

Since $\text{Pred}_2([Fx]_x, [Guv]_{uv}) = [F[Guv]^{uv}]_{uv}$, clause (8.2) insures that $\langle u, v \rangle \in H([F[Guv]^{uv}]_{uv}) \equiv [Guv]^{uv} \in H([Fx]_x)$.

(Here I use, not mention, intensional abstracts from L_ω.)

9. Examples of type 1 and type 2 model structures are easily constructed. E.g., a type 1 model structure can be constructed relative to a model for first-order logic with identity and extensional abstraction, and a type 2 model structure can be constructed relative to a model for first-order logic with identity, extensional abstraction, and Quine's device of corner quotation.
10. $[A(u_1, \ldots, u_p)]_{u_1 \ldots u_p}$ and $[A(v_1, \ldots, v_p)]_{v_1 \ldots v_p}$ are *alphabetic variants iff*$_{\text{df}}$ their externally quantifiable variables are the same and, for each k, $1 \leq k \leq p$, u_k is free in A for v_k and conversely. A term t is said to be free for v_i in A if and only if, for all v_k, if v_k is free in t, then no free occurrence of v_i in A occurs either in a sub-context of the form $(\exists v_k)(\ldots)$ or in a sub-context of the form $[\ldots]_{\alpha v_k \gamma}$. (Recall that $(\forall v_k)(\ldots v_k \ldots)$ is an abbreviation for $\neg(\exists v_k)\neg(\ldots v_k \ldots)$.) Thus, if t is free for v_i in A, the result of substituting t for the free occurrences of v_i in A produces no "collision of variables". Let $A(v_1, \ldots, v_k)$ be any formula; $v_1, \ldots, v_k$ may or may not occur free in A. Then, I write $A(t_1, \ldots, t_k)$ to indicate the formula that results when, for each j, $1 \leq j \leq k$, the term t_j replaces each free occurrence of v_j in A.
11. The notion of t's being free for v_i in A is defined in note 10. Example: if $A(v)$ is $F[Gv]$ and t is $[Hw]$, then $A(t)$ is $F[G[Hw]]$. In this example t is free for v in $A(v)$; v is an externally quantifiable variable in $A(v)$, and w is an externally quantifiable variable in $A(t)$.
12. T1 is the simplest formulation of conception 1. In it the Barcan formula and its converse are derivable. This feature can be removed by slightly complicating the axioms and rules. Corresponding adjustments would then be made in the semantics. For simplicity of exposition these sophistications will not be pursued.

Chapter 3

1. Such a contradiction can be derived using either T1 or T2. Of course, only T2, which is our logic for conception 2, is relevant here, for the paradox of analysis is a puzzle in the logic for intentional matters, to which conception 2 is tailored.
2. See p. 215, Benson Mates, 'Synonymity' (p. 125 in Linsky). I use the term 'Mates' puzzle' to apply to all *prima facie* substitutivity failures involving synonymous predicates (or formulas) such that the substitutivity failure can be traced to some form of ignorance about linguistic, historical, or social matters. For further discussion of Mates' puzzle, see §39.
3. The instance of Mates' puzzle generated by the formula 'x does not know that whatever chews [pronounced *cho͞oz*] chews [pronounced *chôz*]' dramatizes the linguistic character of the problem.
4. I do not count Carnap's analysis of assertion and belief in *Meaning and Necessity* as a serious formal attempt to resolve the paradox of analysis because it is so fraught with problems, including its well-known violation of the Langford-Church translation test. When I speak of Church's resolution of the paradox I refer to a synthesis of the views found in his papers. I do not include the remark in his review of Max Black and Morton White; with tinkering perhaps that remark can be made to mesh with the resolution I offer in this chapter.
5. Church uses this theory of synonymy as the intuitive motivation for the Alternative (0) theory of concepts in his 'Outline', part two. Church holds the principle that the formulas $A(v)$ and $B(v)$ are synonymous if and only if the concept of being a thing v such that $A(v) =$ the concept of being a thing v such that $B(v)$. In this way, his theory of synonymy doubles as theory of concept identity.
6. This is the Leibnizian definition of identity. Frege's and Russell's definitions are substantially the same.
7. In a Churchian language '$y = z$' cannot be converted into '$(\forall f)(fy \equiv fz)$' by rules (1)–(7). The most that can be achieved is the conversion of '$y = z$' into '$(\lambda yz)(\forall f)(fy \equiv fz)(y, z)$' by use of rule (5). And this is not enough to obtain the above instance of the paradox.
8. Pp. 68–71, 'Intensional Isomorphism'.
9. Axiom 66, 'Outline', part two.
10. See p. 69, 'Intensional Isomorphism'.
11. In a Churchian language the rules (1)–(7) at most allow 'Outweighs(y, z)' to be converted into '(λyz)(the weight of y is greater than the weight of z)(y, z)'; they do not permit the production of 'the weight of y is greater than the weight of z'.
12. The ancestor relation is an historically interesting example: x is an ancestor of y *iff* every relation closed under the parent relation holds between x and y; however, 'x is an ancestor of y' and 'every relation closed under the parent relation holds between x and y' are not synonymous isomorphic.
13. The concept of logical consequence for propositions can be defined on analogy with the definition of logical validity for propositions found in §47.
14. And of course what is learned is not a variety of operational facts involving, e.g., ruler and compass; such facts, while not irrelevant, are only incidental.
15. Scope ambiguities provide a helpful analogy: treating them as semantic is inferior to treating them as structural.
16. The following example illustrates the purpose of clause (5). The unanalysed identify concept $[\underline{y = z}]_{yz}$ has two converses, the first one being a concept possessed by people who are sensitive to the fact that it is a converse of $[\underline{y = z}]_{yz}$ and the second one being a concept that can be possessed by people who are not sensitive to this fact. To mark this distinction, I use $[\underline{y = z}]_{zy}$ to denote the

former concept and $\underline{[y = z]_{zy}}$ to denote the latter concept. Generalizing on this, I allow that, for any formula A and any sequence of variables α, the following three intensional abstracts are well-formed: $[A]_\alpha$, $[\underline{A}]_\alpha$, $\underline{[A]_\alpha}$. In the limiting case where $[A]_\alpha$ is a normalized term, $[\underline{A}]_\alpha = \underline{[A]_\alpha}$ is valid; similarly, $[\underline{A}] = \underline{[A]}$ is valid.

17. In this semantics the denotation of the complex term, e.g., $[Gx \mathbin{\&} Hx]_x$ will be the defined concept that is the conjunction of G-ness and H-ness, and the denotation of the complex term $[\underline{Gx \mathbin{\&} Hx}]_x$ will be the undefined concept that is the conjunction of G-ness and H-ness. This suggests an easy solution to the problem of specifying appropriate model structures. Simply build these model structures so that they include two corresponding sorts of conjunction, one that yields defined concepts and one that yields undefined concepts. The same thing goes for each of the other fundamental logical operations, negation, existential generalization, etc. Now type 3 model structures, which are designed to model qualities and complex concepts (see §43), already contain two sorts of conjunction, negation, etc. However, in order to develop the second method mentioned in the text, one would have to construct still another type of model structure (called type $\underline{2}$). A type $\underline{2}$ model structure $\mathscr{M}$ is any structure

$$\begin{aligned}&\langle \mathscr{D}, \mathscr{P}, \mathscr{K}, \mathscr{G}, \mathrm{Id},\\ &\ \mathrm{Conj}^{u}, \mathrm{Neg}^{u}, \mathrm{Exist}^{u}, \mathrm{Exp}^{u}, \mathrm{Inv}^{u}, \mathrm{Conv}^{u}, \mathrm{Ref}^{u}, \mathrm{Pred}^{u}_{0}, \mathrm{Pred}^{u}_{1}, \ldots,\\ &\ \mathrm{Conj}^{d}, \mathrm{Neg}^{d}, \mathrm{Exist}^{d}, \mathrm{Exp}^{d}, \mathrm{Inv}^{d}, \mathrm{Conv}^{d}, \mathrm{Ref}^{d}, \mathrm{Pred}^{d}_{0}, \mathrm{Pred}^{d}_{1}, \ldots\rangle\end{aligned}$$

that simultaneously satisfies the following three conditions. First, the elements of $\mathscr{M}$ are such that the following two diminished structures $\mathscr{M}_{u}$ and $\mathscr{M}_{a}$, respectively, are themselves type 2 model structures:

$$\begin{aligned}&\langle \mathscr{D}, \mathscr{P}, \mathscr{K}, \mathscr{G}, \mathrm{Id},\\ &\ \mathrm{Conj}^{u}, \mathrm{Neg}^{u}, \mathrm{Exist}^{u}, \mathrm{Exp}^{u}, \mathrm{Inv}^{u}, \mathrm{Conv}^{u}, \mathrm{Ref}^{u}, \mathrm{Pred}^{u}_{0}, \mathrm{Pred}^{u}_{1}, \ldots\rangle\end{aligned}$$

$$\begin{aligned}&\langle \mathscr{D}, \mathscr{P}, \mathscr{K}, \mathscr{G}, \mathrm{Id},\\ &\ \mathrm{Conj}^{d}, \mathrm{Neg}^{d}, \mathrm{Exist}^{d}, \mathrm{Exp}^{d}, \mathrm{Inv}^{d}, \mathrm{Conv}^{d}, \mathrm{Ref}^{d}, \mathrm{Pred}^{d}_{0}, \mathrm{Pred}^{d}_{1}, \ldots\rangle.\end{aligned}$$

Secondly, the ranges of the nine sort-u logical operations are disjoint from the ranges of the nine sort-d logical operations. The sort-d operations are to be thought of as those whose values are *defined* concepts, and the sort-u operations are to be thought of as those whose values are undefined concepts. The third condition is a bit cumbersome to state. Put roughly, it is that, for every undefined concept in $\mathscr{D}$, there is an associated concept in $\mathscr{D}$ that is fully defined; and conversely, for every defined concept in $\mathscr{D}$, there is an associated concept in $\mathscr{D}$ that is fully undefined.

18. This general style of resolution can be extended to cover difficult instances of the paradox of analysis such as those produced by analyses of extensional abstraction (§27), number (§32), intensional abstraction (§37), definite descriptions (§38), and so on.

I might mention that there is an entirely different technique for representing the ambiguity in intensional abstracts. For example, suppose that 'the proposition that, for all y, if Fy then Fy' and 'the proposition that, for all y, if Fy then $(Gy \mathbin{\&} Hy)$' denote the same type 2 entity on one reading and different type 2 entities on another. (The second reading is that which pertains to ignorance of definitions.) Then the second reading of 'the proposition that, for all y, if Fy then $(Gy \mathbin{\&} Hy)$' might be represented as follows: $[(\forall y)(Fy \supset y$ has the property that is the conjunction of $[Hz]_z$ and $[Gz]_z)]$. The type 2 intension denoted by this intensional abstract is clearly different from that which is denoted by $[(\forall y)(Fy \supset Fy)]$. Now, using this technique, one could contextually define the intensional abstracts of $\underline{L}_{\omega}$. However, in the text I adopt a primitive underlining technique, rather than this one, since it is readily axiomatizable and since it stears well clear of the logical paradoxes.

Chapter 4

1. I should mention that my remarks on higher-order logic will be aimed at higher-order logic in general. I will not discuss the special case of those second-order logics wherein strings as '$f = g$', '$p = q$', and 'xBp' are ill-formed. Given their inability to express identity (and non-identity) among the very entities over which they quantify, I find these second-order logics both unnatural and non-general. In addition, given their restrictions on the use of propositional variables these second-order logics are of little use in the treatment of modal and intentional matters.
2. The historical reason for selecting Δ to play this role is that, when type subscripts are supplied, Δ is the symbol that plays a somewhat related role in Church's formulation of the logic of sense and denotation. For heuristic purposes it might be helpful to think of the predication relation (the Δ-relation) as the property-theoretic analogue of the ϵ-relation from first-order set theory. However, unlike the ϵ-relation, which includes sets in its range, the predication relation includes properties and relations-in-intension in its range. I reserve judgement here on the Aristotelian question of whether the range of the predication relation includes, in addition to properties and relations-in-intension, objects from other metaphysical categories, e.g., individual particulars, species, quantities, actions, positions, locations, times, stuffs, etc. I do, however, envisage a global logical theory that takes the affirmative on this question. (See my 'Predication and Matter' for a partial defense of this position.) All that is important for the present purposes is that the range of the predication relation should include at least properties and relations. In this connection I also reserve judgement on the Aristotelian question of whether the copula in natural language expresses more than one relation. And, further, I reserve judgement on the Platonic question of whether the copula in natural language is satisfiable in varying degrees, from more perfect to less perfect. What matters for our purposes is that the relation expressed by Δ should include in its extension the relations having-as-a-property and standing-in-a-relation.
3. An attractive alternative first-order representation of this argument can be obtained by adapting Richard Grandy's theory of *anadic logic*. On Grandy's theory predicates need not have any fixed degree. Using this idea, one may treat Δ as a special kind of anadic predicate such that $v_1 \ldots v_n \, \Delta \, w_1 \ldots w_m$ is a well-formed formula, for $n \geq 1$ and $m \geq 0$. Accordingly, the argument in the text could be represented as follows:

$$\frac{xy \, \Delta \, [H^2(x, y)]_{xy} \;\&\; uv \not\Delta \, [H^2(x, y)]_{xy}}{\therefore\; (\exists w)(xy \, \Delta \, w \;\&\; uv \not\Delta \, w)}$$

   Thus, the notation for ordered pair is no longer needed. Although I find this approach to the treatment of predication quite attractive, I will for simplicity pursue the treatment given in the text.
4. For the first-order generalization of ramified type theory, see §26. Incidentally,

not only Russell's but also Church's higher-order logic can be constructed within a first-order theory of PRPs with Δ. Note here the relationship to David Kaplan's 'How to Russell a Frege-Church'. By the way, it is uncertain whether Kaplan's "Russellization" can be made to work for Church's Alternative (0), which is designed for intentional matters.

5. To see this, note that, e.g., $(\forall f)(f=f)$ and $(\forall f)(\forall g)(f=g \vee f \neq g)$ are typical laws of higher-order logic. In these sentences, linguistic predicates occur as linguistic subjects. By the way, I use 'linguistic subject' and 'linguistic predicate' to contrast with 'ontological subject' and 'ontological predicate'. My use comes close to Strawson's use of 'logical subject' and 'logical predicate'; see chapter 8 of his *Individuals*, for example.
6. The subject/predicate distinction plays much the same role in the syntax for first-order intensional languages.
7. Indeed, given the algebraic semantics for L_ω the following principle holds: if x is a k-ary intensional entity expressed by the linguistic predicate F^k_h, then x is denoted by the complex term $[F^k_h(v_1, \ldots, v_k)]_{v_1 \ldots v_k}$.
8. Frege is faced with a dilemma: either he cannot express the sample argument, or he too is committed to holding that red and blue are functions.
9. Some might expect that, since this first-order theory for the predication relation contains existence assertions, it is not really logic. (See pp. 195–7, Parsons, 'Frege's Theory of Number'.) In §36 I will argue against this point of view. For the present suffice it to say that even in the case of language L_ω without Δ all standard models are infinite and, indeed, they include denotable properties whose extensions include ω-sequences.
10. Whether there is a complete axiomatization for the valid$_1$ (valid$_2$) sentences of the first-order extensional language L (see §37) is an open question. If there is not, the chart—and, hence, the thesis of the present section—would need to be complicated accordingly.
11. See pp. 195–7, Parsons, 'Frege's Theory of Number' for a defense of this view.
12. Leśniewski stands as one of the few exceptions.
13. Henkin's quasi-completeness result for higher-order quantification theory provides another point of view on the question of the origin of incompleteness in higher-order theories. The relationship between these two points of view is a topic for further reflection.
14. When 'fx' is read aloud in English, we say, 'x is f'!
15. This rule and the Ref-rule are from the definition of what a model structure is; see p. 51 f.
16. The methodological discussion in §37 is relevant to the issue of whether classical logic should be tampered with here.
17. The same thing can easily be done for Quine's resolutions and for the more recent Fitch-Gilmore-Feferman resolution. (An idea analogous to Fitch's original insight lies behind Kripke's recent resolution of the Epimenides paradox in 'Outline of a Theory of Truth'.) For adaptation to the logic for L_ω with predication, Gilmore's lucid paper 'The Consistency of Partial Set Theory Without Extensionality' is ideal.
18. Strictly speaking this principle of predication is fashioned after the class-building principle in the Kelly-Morse set theory (appendix, Kelly, *General Topology*) rather than the von Neumann-Gödel-Bernays set theory. A strict von Neumann-style principle of predication is obtained by restricting the range of all quantified variables in the formula A to safe objects. See note 26.
19. In the setting of T2′ special care is required in formulating the power axioms. See note 27. By the way, a nice feature of the ZF-style theory is that in it = can be defined in terms of Δ: $x = y$ *iff*$_{\mathrm{df}}$ $(\forall z)(x \,\Delta\, z \equiv y \,\Delta\, z)$.

20. The ZF-style theory in the setting of T1 is consistent if ZF in the setting of first-order logic with extensional abstraction (see §15) is consistent. And the GB-style theory in the setting of T1 is consistent if Kelly-Morse set theory in the setting of first-order logic with extensional abstraction is consistent. But the relative consistency of the ZF and GB-style theories in the setting of T2′ remains to be studied.

By the way, it is widely believed that Frege's law V, i.e., $\{x: f(x)\} = \{x: g(x)\} \equiv (\forall x)(f(x) \equiv g(x))$, is responsible for the logical paradoxes in Frege's logic. True enough, the system does appear to be free of these paradoxes if law V is dropped. However, the sytem is also free of them if law V is retained and, instead, the higher-order variables and quantifiers are dropped, thereby converting the system into a first-order logic. The resulting system is virtually the same as first-order logic with extensional abstraction, which we know to be sound and complete. In that logic the first-order counterpart of law V, i.e., $\{x: A\} = \{x: B\} \equiv (\forall x)(A \equiv B)$, is just one of the axioms. This goes to show that Frege's higher-order syntax is as responsible for the logical paradoxes in his logic as is law V. Law V generates no such paradox if, as I have urged in the present chapter, first-order syntax is taken as canonical.

21. As indicated in §22, the following is an example of a definition of truth tailored to T2′:

$$Tx \textit{ iff}_{\mathrm{df}} (\exists y)(x \approx_{\mathrm{N}} [(\exists z)z \,\Delta\, y]^{y} \,\&\, (\exists z)z \,\Delta\, y).$$

And the following is an example of a definition of truth tailored to T1:

$$Tx \textit{ iff}_{\mathrm{df}} (\exists y)(x = [(\exists z)z \,\Delta\, y]^{y} \,\&\, (\exists z)z \,\Delta\, y).$$

The proofs that $T[A] \equiv A$ are straightforward proofs using pairing, abstraction, and null axioms, and $\mathscr{A}17$ or A8. It should be stressed, however, that neither of the above definitions is the philosophically motivated definition based on the correspondence theory of truth given in §45.

22. Corner quotes can be eliminated in favor of Gödel numbers, which can be defined in terms of Δ.

23. See footnote 25, p. 758, Church, 'Comparison of Russell's Resolution of the Semantical Antinomies with that of Tarski'.

24. So if T is defined as in the third line of note 21, then this occurrence of v_i cannot be bound by $(\exists v_i)$ in a context of the form

$$(\exists v_i)(v_k \approx_{\mathrm{N}} [(\exists v_j)v_j \,\Delta\, v_i]^{v_i} \,\&\, (\exists v_j)v_j \,\Delta\, v_i)$$

or

$$(\exists v_j)(v_k \approx_{\mathrm{N}} [(\exists v_i)v_i \,\Delta\, v_j]^{v_j} \,\&\, (\exists v_i)v_i \,\Delta\, v_j).$$

25. For remarks relevant to this distinction between predicative and impredicative principles of predication, see pp. 52–4, Chihara, *Ontology and the Vicious-Circle Principle*, and pp. 758–60, Church, 'Comparison'.

26. This predicative principle of predication is the intensional analogue of the von Neumann, as opposed to the Kelly-Morse, class-building principle. See note 18.

27. Special adjustments are needed in the counterparts of the ZF and GB power principles:

(ZF-Style Power Principle)
$$(\exists y)(\forall x)((\exists z)(z \doteqdot x \,\&\, z \,\Delta\, y) \equiv x \subseteq w)$$

(GB-Style Power Principle)
$$(\exists b)(\forall a)((\exists c)(c \doteqdot a \,\&\, c \,\Delta\, b) \equiv a \subseteq d)$$

where $\alpha \doteqdot \beta$ *iff*$_{\text{df}}$ $(\forall\delta)(\delta \mathbin{\Delta} \alpha \equiv \delta \mathbin{\Delta} \beta)$ and $\alpha \subseteq \beta$ *iff*$_{\text{df}}$ $(\forall\delta)(\delta \mathbin{\Delta} \alpha \supset \delta \mathbin{\Delta} \beta)$. Special adjustments are also needed in the counterparts of the ZF replacement axiom as follows:

(ZF-Style Replacement Principle)

$$(\forall x, y, z)((A(x, y) \mathbin{\&} A(x, z)) \supset y = z) \supset (\exists w)(\forall y)((\exists z)(z \doteqdot y \mathbin{\&} z \mathbin{\Delta} w) \equiv (\exists x)(x \mathbin{\Delta} u \mathbin{\&} A(x, y)))$$

where w is distinct from y and is not free in A. (This principle could, if necessary, be restricted further by permitting no predicates beyond Δ and $\doteqdot$ to occur in A.) The GB-style replacement principle can be formulated on analogy with the ZF-style replacement principle.

28. On the semantical theory defended in §38 there is only one fundamental kind of semantical relation—namely, meaning—and all other semantical relations are derivative, being definable in terms of this one kind of meaning. If this theory is right, there is no reason to treat explicitly any semantical paradoxes beyond those generated by this one kind of meaning.

There is also no need to impose the indicated restrictions concerning grounded formulas in the setting of conception 1. The reason is that both meaning and intentionality are conception 2 phenomena. (The thesis that meaning is a conception 2 phenomenon is defended in §38. This thesis, coupled with the thesis that there is just one underlying kind of meaning, frees us to construct unrestricted ZF and GB-style theories of conception 1 PRPs. These conception 1 theories make it possible to give an especially simple no-class construction of pure ZF and GB set theory (see §31), and this in turn simplifies the argument of §§30–2. But the conclusion of that argument can also be won in the setting of conception 2; see note 16 in the next chapter for an indication of how to do this by means of an attractive informal argument.)

In comparison with Tarski-style and Russell-style resolutions of the semantical paradoxes, the proposed resolution permits a language to have a single univocal meaning predicate and single univocal truth predicate rather than an infinite hierarchy of meaning and truth predicates. Such infinite hierarchies of meaning and truth predicates lead to violations of Davidson's finite learnability requirement (see desideratum 13 in §4). And in comparison with a resolution of the semantical paradoxes that is fashioned after Kripke's technique in 'Outline', the resolution proposed here permits a meaning predicate M^2 that suffers from no "gaps": $M^2(\ulcorner A(\alpha)\urcorner, [A(\alpha)]_\alpha)$ holds for all L_ω-formulas $A(\alpha)$ even where M^2 is itself a predicate in $A(\alpha)$. A further advantage this resolution has over those given in the style of Tarski or Kripke is that it does not regard the semantical and intentional paradoxes as unrelated phenomena; rather it resolves them both by one and the same account.

If certain arguable modal principles about meaning (e.g., x means $y \supset \Box x$ means y) were adjoined to the ZF and GB-style theories described in the text, then certain new semantical paradoxes could be derived. But this threat evidently can be shortcircuited by restricting T2′ axiom $\mathscr{A}$17 to grounded formulas A_u and B_u: $\Box(A_u \equiv_\alpha B_u) \equiv [A_u]_\alpha \approx_{\mathrm{N}} [B_u]_\alpha$. (Analogous restrictions could, if necessary, be imposed elsewhere in T2′.) Even with this restriction, the schema $T[A_u] \equiv A_u$ can still be derived for the univocal truth predicate T which is definable in the theory.

29. A somewhat related pragmatic resolution of the semantical paradoxes is suggested in Charles Parsons' 'The Liar Paradox'. Tyler Burge also argues for a pragmatic resolution in 'Semantical Paradox'; however, he locates the pragmatic element in an indexical slot in a truth predicate 'true$_i$' for sentences. In my tentative resolution the truth predicate for propositions is treated as a 1-place

predicate definable in terms of Δ and $\approx_N$; the pragmatic element is instead located in the contextually determined implicit universe of discourse u. Although Burge criticizes (footnote 13, p. 176) Parsons' shift-in-the-domain-of-discourse resolution, Burge's resolution may be viewed as a derivative form of the version of the shift-in-the-domain-of-discourse resolution that I have described. To see this, notice that a truth predicate 'true$_u$' for sentences, where u is an indexical slot, is definable as follows: true$_u(x)$ *iff*$_{df}$ $(\exists y)(y \,\Delta\, u \,\&\, M^2(x, y) \,\&\, Ty)$. Here u limits not the universe of *sentences* but rather the universe of *propositions*.

30. For a discussion of an analogous relationship between ordinary ZF and simple type theory, see pp. 266–86, Quine, *Set Theory and Its Logic*, revised edition.
31. Let '$A(x, y)$' abbreviate:

    formula x expresses set y & $\langle x, y\rangle \in \{xy\colon x$ is a formula & $y = \emptyset\}$.

    And let '"$A(x, y)$" expresses the set $\{xy\colon A(x, y)\}$' be adopted as an axiom. (Assume that 'is a formula' has been defined in terms of Gödel numbers, which in turn are defined in terms of $\in$; assume also that $\{xy\colon \ldots\}$ is contextually defined in terms of $\in$. Then one can easily derive contradictions in ZF and GB set theories even when their comprehension schemas $(\exists y)(\forall x)(x \in y \equiv (x \in u \,\&\, B(x)))$ and $(\exists y)(\forall a)(a \in y \equiv B(a))$ are restricted to "grounded" set-theoretical formulas B (i.e., set-theoretical formulas B whose quantified variables are all restricted in their ranges to antecedently given sets). No analogous contradiction is derivable in T2′ from the proposed ZF-style or GB-style axioms for Δ. The reason is that there is no principle of extensionality in these logics for intensional entities.

Chapter 5

1. Later in this section I give an example of how the idea of set might be "genetically" related to ideas of certain naturalistic objects. Incidentally, in the present discussion I do not assume that the naturalistic ontology of packs, bunches, flocks, etc. is justified. The point rather is that, if this ontology is justified, that would confer no justification on the ontology of sets since sets are quite unlike packs, bunches, flocks, etc. I should also mention that since writing this section I have learned that Ruth Barcan Marcus makes many similar points in her 'Classes, Collections, and Individuals'.
2. For another sort of problem, suppose that by time t some given bunch of grapes has dwindled down to a single grape. Does the bunch still exist? If so, is the bunch = the grape? I'm not sure. But notice that in set theory the answers are already prescribed: the singleton of the grape exists, and it is not the same thing as the grape. How bizarre singleton sets and null sets are.
3. This difference gives rise to another: the time invariant principle of extensionality, which is supposed to be valid for the sets of set theory, is not valid for ordinary collections, social classes, and ordinary sets. Consider art collections. It is in principle possible that the Tate collection should contain at t exactly those art works that the Guggenheim collection contains at $t'(t \neq t')$ and yet that the Tate collection and the Guggenheim collection should always remain distinct.
4. It might be objected that what I say in the text in this and in the following paragraph results from a confusion between membership ($\in$) and inclusion ($\subseteq$). However, this objection begs the question. The alleged membership/inclusion distinction is a set-theoretical distinction, yet what is presently at issue is whether set-theoretical concepts can be justified by appealing to the ontology of

ordinary collections, social classes, and ordinary sets. In controversies such as this, one has no choice but to fall back on naive intuition, and naive intuition concerning the relation of being in ordinary collections, social classes, and ordinary sets provides *prima facie* evidence for my conclusions.

By the way, doubts about the $\in$/$\in$-inclusion distinction in set theory do not carry over to the Δ/Δ-inclusion distinction in the theory of PRPs.

5. What is the aggregate of things that are not in themselves?
6. See, e.g., Tarski, 'On the Geometry of Solids'.
7. Not unrelatedly, Gödel attempts to justify set theory on grounds of intuitions about what he calls "pluralities", pp. 137 ff., 'Russell's Mathematical Logic' (pp. 220 ff. in Benacerraf and Putnam).
8. I am inclined to treat plurals in this way. For example, the second problematical sentence cited a moment earlier in the text might be provisionally thought of as follows:

   ($\exists x, y$)(x = (being a) county & y = (being a) state & the-x-aggregate occupies the same territory as the-y-aggregate & the-x-extension outnumbers the-y-extension & the-typical-x resents federal intervention more than the-typical-y).

   The expressions ⌜the-α-aggregate⌝, ⌜the-α-extension⌝, and ⌜the-typical-α⌝ can then be treated as contextually defined operators on α. For example, if ⌜the-α-extension⌝ is represented by ⌜$\{\beta : \beta \,\Delta\, \alpha\}$⌝, it can be contextually defined in the way suggested in the text a bit later. Incidentally, it might be crucial that α ranges over properties rather than sets, given the extensionality of sets. Is it not true that the typical policeman $\neq$ the typical short-order cook (or at least that the ideal policeman $\neq$ the ideal short-order cook) and that this would be so even if, because of widespread moonlighting practices, all and only policemen coincidentally turned out to be short-order cooks?
9. Note, this is just the first-order version of Frege's law V. See note 20, chapter 4.
10. This is of course part of Russell's theory of meaning. Russell's theory is unlike the Fregean theory according to which, not only do predicates and formulas express something, but also they name something. I defend Russell's theory over Frege's in §38.
11. This is just the first-order analogue of Russell's higher-order "no-class" definition of extensional abstracts. So it is clear that the Russellian theory of meaning and the no-class theory are of a piece. Incidentally, even Quine acknowledges, 'Classes may be thought of as properties in abstraction from any differences which are not reflected in differences of instances' (pp. 120–1, *Mathematical Logic*, revised edition).
12. Question: on this account what is the primary semantical correlate of the extensional abstract $\{v_i : Av_i\}$? Answer: nothing in particular because on this account the extensional abstract $\{v_i : Av_i\}$ is only an *indefinite description* of a property (specifically, a property that is co-extensive with the property expressed by the formula Av_i).
13. I speak of extensional semantics in the sense of Tarski, Carnap, and their followers.
14. Or *ideal*, e.g., the typical county, the typical state, the ideal policeman, etc. See note 8.
15. These no-class constructions entail that ZF (GB) with extensionality is consistent if ZF (GB) without extensionality is consistent. These relative consistency results differ from those obtained by other authors. For more on this topic, see notes 23 and 26. By the way, the no-class construction of GB can be adapted to obtain a no-class construction of—and a relative consistency result for—Kelly-Morse set theory.

16. There is a more direct, though less rigorous, way to win this conclusion. (See my 'Foundations Without Sets' for elaboration of this line of argument.) One need only show that pure and applied set theory can be interpreted informally as theories of properties. This can be done for the pure set theories ZF and GB by informally interpreting them as theories of an appropriate kind of property, for example, pure L-determinate type 1 properties. (Property x is L-determinate *iff* $\Box(\forall y)(y \Delta x \supset \Box y \Delta x)$, and x is pure L-determinate *iff* x is L-determinate, the instances of x are L-determinate, their instances are L-determinate, and so on all the way down.) Since type 1 properties are identical if necessarily equivalent, pure L-determinate properties will be identical if they have the same pure L-determinate instances. But this is just what the axiom of extensionality says when pure set theory is interpreted as a theory of pure L-determinate properties, so a universe of pure L-determinate properties validates this axiom. (Pure set theory can also be interpreted as a theory of a special kind of conception 2 property; see note 18 for an example.) Applied set theory, on the other hand, can be interpreted as a theory of properties in which identity for empirical sets is read simply as equivalence for empirical properties. When the principle of extensionality for empirical sets is interpreted this way, it is a trivial tautology. It will become clear that the no-class constructions in the text are simply formalizations of these informal interpretations derived within ZF and GB-style logics for the predication relation.
17. See George Boolos, 'The Iterative Conception of Set'.
18. Someone might doubt that the property $[y$ is an L-determinate property whose instances are instances of a property formed at a stage prior to $\alpha]_y$ has "enough" instances to validate the relevant Zermelo-style axioms. This doubt is unfounded, however, for this property is necessary equivalent to:

    $[(\exists u)(u$ is an aggregate of properties that are instances of a property formed at a stage prior to α & $y = [v$ is a property in $u]_v)]_y$.

    Here I use the notion of *aggregate* which was characterized in the previous section. (The notion of *sum* from an unrestricted part/whole logic would serve our purposes equally well.) By using the notion of aggregate (sum), one obtains a property that clearly has "enough" instances to validate the relevant Zermelo-style axioms. At the same time, one avoids the sort of circularity found in the set-theoretical motivation that Russell gave for the axioms of reducibility in the setting of his no-class theory (see pp. 80–3, 'Mathematical Logic as Based on the Theory of Types', *Logic and Knowledge*).

    By the way, if one were to use the formulation given in this note, one could also construct an iterative hierarchy of type 2 properties that validates a ZF-style theory for conception 2. Such a theory would be especially congenial with the picture of concepts given in chapter 8.

    Since Platonists object to the idea that PRPs, type 1 or type 2, are really "formed", one might better think of these hierarchies as stage-by-stage certifications of sub-portions of the extension of the predication relation over the field of PRPs.
19. Here $\subseteq$ is defined in terms of Δ: $u \subseteq v$ *iff*$_{\mathrm{df}}$ $(\forall w)(w \Delta u \supset w \Delta v)$. Recall from §22 that in T1 x is a property *iff*$_{\mathrm{df}}$ $(\exists y)x = [z \Delta y]^y_z$ and from §29 that x is L-determinate *iff*$_{\mathrm{df}}$ $\Box(\forall y)(y \Delta x \supset \Box y \Delta x)$.
20. Sentence A may be taken from ZF or GB. If it is taken from GB, then I will assume that its special set variables $a, b, c, \ldots$ are contextually defined in terms of $\in$.

21. (Comprehension) $x \Delta [x \Delta u \,\&\, A]^{u}_{x} \equiv (x \Delta u \,\&\, A)$
(Null) $x \Delta [x \neq x]_x \equiv x \neq x$
(Pairing) $x \Delta [x = u \vee x = v]^{uv}_{x} \equiv (x = u \vee x = v)$
(Union) $x \Delta [(\exists z)(x \Delta z \,\&\, z \Delta u)]^{u}_{x} \equiv (\exists z)(x \Delta z \,\&\, z \Delta u)$
(Power) $x \Delta [(\forall z)(z \Delta x \supset z \Delta u)]^{u}_{x} \equiv (\forall z)(z \Delta x \supset z \Delta u)$
(Infinity) $(\exists v)([x \neq x]_x \Delta v \,\&\, (\forall z)(z \Delta v \supset [w \Delta z \vee w = z]^{z}_{w} \Delta v))$
(Replacement) $(\forall xyz)((A(x, y) \,\&\, A(x, z)) \supset y = z) \supset$
$(y \Delta [(\exists x)(x \Delta u \,\&\, A(x, y))]^{u}_{y} \equiv_y (\exists x)(x \Delta u \,\&\, A(x, y)))$
(Regularity) $(\forall x)((\exists y)y \Delta x \supset (\exists y)(y \Delta x \,\&\, (\forall z)(z \Delta x \supset z \not\Delta y)))$.
These axioms do not conflict with the resolution of the intentional and semantical paradoxes given in §26. For I am presently working in the setting of conception 1 whereas the intentional and semantical paradoxes—and their resolution—fall within conception 2. If conception 2 entities were brought in directly or indirectly, qualifications would be in order.

22. T1 rules R1–R3 may be applied directly to any axiom in TZF$^-$. Note that TZF$^-$ is consistent if ZF in the setting of first-order logic with extensional abstraction (see §15) is consistent.

23. A common way to give a consistency proof for one theory relative to another is to model the first theory within the second. This technique is frequently used to prove that set theory is consistent if a given sub-theory with fewer axioms is consistent. We have seen that the relevant difference between set theory and the logic for the predication relation lies in the presence or absence of the axiom of extensionality. So the logic for the predication relation will model set theory if set theory with the axiom of extensionality can be proved in this way to be consistent relative to set theory without that axiom. In 'More on the Axiom of Extensionality' Dana Scott shows that ZF with extensionality cannot be proved consistent relative to ZF without extensionality, at least when no abstraction operation is taken as primitive. The present result has the force of proving the desired relative consistency for ZF when an abstraction operation—either intensional or extensional—is taken as primitive. By the way, if modifications in the statement of the original ZF axioms are permitted, ZF with extensionality can be proved consistent relative to ZF without extensionality. For example, Scott reports this for a modified (but equivalent) formulation of ZF in which $(\forall v)(v \in y \equiv v \in z)$ replaces $y = z$ in the antecedent of the replacement axiom. Indeed, ZF can be proved consistent relative to a certain modified ZF-style intuitionistic set theory that lacks the axiom of extensionality (Harvey Friedman, 'The Consistency of Classical Set Theory Relative to a Set Theory with Intuitionistic Logic'). All these relative-consistency results may double as no-class constructions of ZF.

24. (Comprehension) $a \Delta [A(v)]_v \equiv A(a)$
(Null) $(\exists a)(\forall b)(b \Delta a \equiv b \neq b)$
(Pairing) $(\exists a)(\forall b)(b \Delta a \equiv (b = c \vee b = d))$
(Union) $(\exists a)(\forall b)(b \Delta a \equiv (\exists c)(b \Delta c \,\&\, c \Delta d))$
(Power) $(\exists a)(\forall x)(x \Delta a \equiv x \subseteq b)$
(Infinity) $(\exists a)([x \neq x]_x \Delta a \,\&\, (\forall b)(b \Delta a \supset [y \Delta b \vee y = b]^{b}_{y} \Delta a))$
(Replacement) $(\forall bcd)((\langle b, c\rangle \Delta x \,\&\, \langle b, d\rangle \Delta x) \supset c = d) \supset$
$(\exists a)(\forall c)(c \Delta a \equiv (\exists b)(b \Delta e \,\&\, \langle b, c\rangle \Delta x))$
(Regularity) $(\forall x)((\exists a)a \Delta x \supset (\exists a)(a \Delta x \,\&\, (\forall b)(b \Delta x \supset b \not\Delta a)))$.
Here $a, b, c, \ldots$ are contextually defined variables ranging over safe entities, i.e., entities that have properties; all quantified variables in the formula A are restricted in their range to safe entities, and v is free for a in A and conversely. The supplementary remarks in note 21 also apply *mutatis mutandis* to these GB-style axioms.

25. T1 rules R1–R3 may be applied directly to any axiom in TGB^-. Note that TGB^- is consistent if GB in the setting of first-order logic with extensional abstraction (see §15) is consistent.
26. One might think that a no-class construction of GB is already at hand due to a relative consistency result of R. O. Gandy ('On the Axiom of Extensionality', part two). Gandy showed that a certain modified version of von Neumann-Gödel-Bernays class theory with extensionality is consistent relative to that theory without extensionality. This result does not provide what we need, though. Gandy's modified version of GB contains a primitive extensional abstraction operator (λv_i) for which $(\lambda v_i)A = (\lambda v_i)B \equiv (\forall v_i)(A \equiv B)$ holds whenever $(\lambda v_i)A$ and $(\lambda v_i)B$ are well-formed. Thus, (λv_i) behaves like the primitive class-abstraction operator $\{v_i\colon \ldots\}$. By adapting Dana Scott's result (see note 23), one evidently can show that Gandy's proof fails if the abstraction operator (λv_i) is not taken as primitive but instead is contextually defined in terms of ϵ on analogy with the contextual definition of $\{x\colon \ldots\}$ that I suggested in §28. But what we need is a theory lacking extensionality which can model class theory with extensionality and in which all extensional abstraction operators can be contextually defined. So Gandy's result seems not to do the job. The intensional abstraction operation of L_ω is what is wanted in the no-class construction of GB.
27. A similar no-class construction is also possible for a GB-style applied set theory. And no-class constructions for pure and applied set theory are possible in the setting of conception 2.

Chapter 6

1. In place of (5) he might say, equivalently, that for all properties z of natural numbers, if 0 has z and the successor of each natural number having z itself has z, then every natural number has z.
2. Recall that the first-order intensional logic T2—unlike, say, first-order quantifier logic—is committed to an infinite ontology of PRPs. By the way, the neo-Fregean definitions work in the setting of the logic T1. However, in order to derive Peano's postulates in T1, logical principles for Δ must be adjoined. I do not discuss the T1 approach in the text only because the T2 approach is so neat.
3. This is not true for Russell, according to whom sets do not really exist; Russell's entities are propositional functions. This is not just a scholarly point, for with a little fiddling Russell can easily avoid the criticism.
4. Note that 'are' is the plural form of the copula 'is'.
5. Recall that since the extensional abstract $\{v\colon Av\}$ is contextually defined in terms of Δ, its use carries ontological commitment to properties, not to sets.

   I emphasize, however, that this treatment of extensional abstracts is only tentative. One attractive alternative is to treat extensional abstracts as denoting ordinary aggregates. Someone who adopts this alternative would be led naturally to the sort of conclusion reached in Glenn Kessler's 'Frege, Mill, and the Foundations of Arithmetic', namely, that numbers are relations-in-intension holding between ordinary aggregates and properties (where the role of these properties is to provide a principle by which to identify—and hence, to count—things in the aggregates). This, though, is only a slight variation on the logicist position I am defending. Numbers still would be intensional entities, and if the "aggregate slot" in a relation that is a number is treated as a certain kind of parameter, then number theory still can be construed as part of pure intensional logic with Δ.

6. Constructions involving 'the number of *F*s', 'as many as', 'more than', 'less than', and other verbal forms from the idiom of cardinality can be easily defined by neo-Fregean means in L_ω using contextually defined definite descriptions, extensional abstracts, and Δ. Implicit in these definitions is Frege's well-argued, rather uncontroversial thesis that the natural numbers are cardinal numbers as opposed to ordinal numbers or quantities in the sense of amounts. To my knowledge there are no good arguments for either of these opposing views.
7. See pp. 57–8, Paul Benacerraf, 'Numbers'.
8. Near the end of the paper Benacerraf makes a positive proposal about how to analyse number-theoretic language, not in terms of particular objects called numbers, but *indefinitely* in terms of whole structures that behave in the appropriate way. But if the argument just summarized in the text were valid, then evidently a variant of it would apply against this positive proposal:

   There are many different ways—e.g., Frege's original way, the neo-Fregean way, Benacerraf's way, etc.—in which, for all we know, number-theoretic language could be correctly analysed.

   ∴ Number-theoretic language could not be correctly analysed in any of these ways.

9. This fact is what guided Frege's original research into the analysis of the natural numbers. And it is evidently one that Paul Benacerraf would accept, for he endorses '...the concern for having a homogeneous semantical theory in which semantics for the propositions of mathematics parallels the semantics for the rest of language'. (p. 661, Benacerraf, 'Mathematical Truth')
10. I should mention that Benacerraf does consider sentences of the form 'The *F*s are *n*' on pp. 58–60, where he says that, e.g., 'The lions in the zoo are seventeen' probably comes into the language by deletion from 'The lions in the zoo are seventeen in number', which in turn probably derives from something like 'Seventeen lions are in the zoo'. However, in view of the fact that the problem raised by Benacerraf's criticism is a species of the indeterminacy problem in logico-linguistic theory, one wonders whether it is consistent to use these assertions in an argument against logicism. (See note 8 above.) Waiving this reservation, however, one would like to know what is the logical form of 'Seventeen lions are in the zoo' and what is the analysis of 'seventeen' as it occurs in this sentence. The answer should permit an account of the inference from 'There are seventeen lions in the zoo' and 'Fifteen plus two is seventeen' to 'There are fifteen plus two lions in the zoo' and also an account of the equinumerosity principle stated earlier in this section. This indicates that, even if numerical adjectives in natural language were operators, as Benacerraf suggests (p. 60), they must nonetheless have semantical correlates that behave with respect to each other in exactly the same way that the natural numbers do. This seems to be reason enough for identifying the semantical correlates of numerical adjectives with the natural numbers themselves. In any event, does not the "arithmetic" for the semantical correlates of numerical adjectives fall squarely within the province of natural logic, and is this not all the logicist needs to make good his philosophy? If so, why not eschew the operator approach to numerical adjectives and return to the essentially simpler neo-Fregean theory?

    Incidentally, if one thinks a bit about the interesting grammatical phenomena cited by Benacerraf on p. 60 (center), one sees that they can be nicely predicted by the neo-Fregean theory.
11. A complete set of axioms for $\langle NN, =, 0, ' \rangle$ consists of (1′)–(4′) plus the following: (where $\bar{n}$ stands for *n* consecutive occurrences of ′)

(6) $(\forall x)(NNx \supset (x \neq 0 \supset (\exists y)(NNy \,\&\, x = y')))$
(7.1) $(\forall x)x \neq x'$
(7.2) $(\forall x)x \neq x''$
$\vdots$
(7.*n*) $(\forall x)x \neq x^n$
$\vdots$

See, e.g., §3.1 'Natural Numbers with Successor' in Enderton, *A Mathematical Introduction to Logic* and Quine's comments on axioms (7.*n*), p. 99 in *From a Logical Point of View*. Axiom (6) follows from (1′), (2′), (5′), plus the validity under discussion in the text. Axioms (7.*n*) follow from axiom $\mathscr{A}$12.

12. See, e.g., Landau, *Foundations of Analysis*, for a set-theoretic construction of number theory with + and · and of real and complex analysis beginning with just Peano's postulates.
13. Indeed, identity and necessary equivalence are definable in terms of predication. For the definition of =, see note 19, chapter 4. For the definition of $\approx_N$, see §46.
14. On this point, the logicism I am defending is possibly closer to Russell's than Frege's, for Russell makes it clear that his logicism requires only that the truths of mathematics be logical validities.
15. Question: in a Zermelo-style theory how many singleton properties are Δ-instances of the property with which the number 1 has been identified on my neo-Fregean analysis? Answer: as many as you like up to unsafe points.
16. Strictly speaking, the Kelly-Morse style theory (see appendix, Kelly, *General Topology*).
17. Allusions to this view can be found in Gödel's paper 'Russell's Mathematical Logic', pp. 137 ff. (pp. 220 ff. in Benacerraf and Putnam).
18. On Hilbert's view, like Frege's view before it, proof is the key to the account of our knowledge of complex logical truths. I am doubtful that the role of proof is as great as Hilbert and Frege thought. But such doubts do not affect the argument I give in the text.
19. On the justification of the laws of arithmetic (as opposed to analysis and set theory), Hilbert himself would give a reply that falls back on intuitions.
20. For more on this see p. 75 ff., Chihara, *Ontology*, and p. 674, Benacerraf, 'Mathematical Truth'.

Chapter 7

1. In *Meaning and Necessity* Carnap does not take this attitude toward all intensionality in language. In particular, he gives a fully extensional ("formal-mode") account of 'belief'-sentences. This account is a descendant of the account given in his *The Logical Syntax of Language*, which in turn appears to have been derived from the account given in Wittgenstein's *Tractatus*. The fact that in *Meaning and Necessity* Carnap offers no unified account of all intensionality in language would seem to be a count against his theory.
2. Russell had a rather complicated theory of extensionality and intensionality in language. For Russell there did exist certain *prima facie* cases of intensionality and extensionality that were not at all what they seemed to be. For example, all *prima facie* violations of Leibniz's law were deemed only apparent and were explained away by means of the theory of descriptions. In a similar fashion, Russell also explained away the *prima facie* extensionality generated by exten-

sional abstracts. This he did by means of his no-class analysis. But Russell held that there existed some genuine intensionality in language, for he took at face value all the usual *prima facie* violations of the principle of the substitutivity of equivalents. At the same time, he held that there is some genuine extensionality in language. Specifically, he held that Leibniz's law is universally valid. And, in the same vein, he held that the logical connectives are extensional in the sense that they are truth-functional and, hence, that the principle of the substitutivity of equivalents holds for all contexts that are exclusively built up by means of them. Thus, in the final analysis Russell subscribed to the view that language does bifurcate into two ultimate kinds, intensional and extensional. In this, his view is on a par with the views of C. I. Lewis, the Carnap of *Meaning and Necessity*, Hintikka, Montague, Kripke, *et al.*

3. In linguisitc theory the conflict between interpretive and generative semantics is somewhat similar in flavor.
4. Strictly speaking, definite descriptions are formed by applying an unanalysed term-forming operator to what Frege calls concept-names. However, this fine point is immaterial here. Incidentally, in the text I confine my comments to Frege's theory of definite descriptions in natural language. His treatment of definite descriptions in the formal language of the *Grundgesetze* is different; a special variant of the "favorite-object" approach is taken there on the matter of vacuous descriptions.
5. Analogously, it appears that in 'Quantifying In' David Kaplan praises Frege's theory of intensionality from the conservative point of view while in 'What is Russell's Theory of Descriptions?' he criticizes Russell's theory of descriptions from the liberal point of view. At many points in contemporary work in modal logic there also appear to be methodological vacillations over these issues.
6. See also note 14, chapter 1.
7. Apparent violations of Leibniz's law produced by extensional abstracts and functional constants can be explained away by analogous means. By the way, Russell's theory of descriptions is not essential to the program in the text. It would be possible, though more complex, to treat definite descriptions much as Frege does. However, that would force me to enrich my algebraic model structures with appropriate new logical operations to handle definite-description concepts (and, then, to make adjustments for certain new propositions that would lack a truth value).
8. The intensional language $\underset{\sim}{\mathrm{L}}_{\omega}$, introduced in §20, can be translated into a finite-based first-order extensional language in much the same way as L_{ω}.
9. There are only heuristic reasons for choosing the symbols $\bar{F}_i^j$ for the names I need; any arbitrary symbol would suffice as long as one gives L the prescribed interpretation. Incidentally, when one interprets L as intended, no instances of Frege's '$a = a$'/'$a = b$' puzzle arise. (Frege's puzzle is discussed in the following section.)

In what follows I will show how contextually to define intensional abstracts that contain predicates selected from $F_1^1, \ldots, F_p^q$. One could extend this method to intensional abstracts that contain, in addition, predicates selected from $\mathscr{Conj}^3, \ldots, \mathscr{Pred}^3$. One would adjoin $\overline{\mathscr{Conj}}, \ldots, \overline{\mathscr{Pred}}$ to the primitive names of L. This step calls for complications in the specification of the semantics of L, for in this region one treads close to certain logical paradoxes. I should mention also, that there exist simplifications of the construction given in the text. For example, all predicates beyond Δ might prove to be definable in terms of appropriate primitive names plus Δ. The construction, therefore, should not be thought of as definitive; I give it in order to show that the thesis of extensionality is defensible.

10. $\mathscr{M}$ is a model structure; $\mathscr{I}$ is an intensional interpretation of the primitive predicates of L_ω; $\mathscr{D}_{\mathscr{M}}$ is the domain of discourse in $\mathscr{M}$; Conj, Neg, ... are the fundamental algebraic logical operations in $\mathscr{M}$; $\bigcup_{k \geq 0} \text{Pred}_k$ is the union of the predication operations $\text{Pred}_0, \text{Pred}_1, \ldots$ in $\mathscr{M}$, and $\mathscr{G}$ is the "actual-extension" function in $\mathscr{M}$.
11. The following corollary is an immediate consequence: if L_ω-formula A translates into L-formula A^*, then A is valid$_1$ (valid$_2$) *iff* A^* is valid$_1$ (valid$_2$). In view of this, T1 and T2 may be viewed as logics for L. This raises the question of whether there is a complete recursive axiomatization of the valid$_1$ or valid$_2$ formulas of L. If there is not, it is all but certain that the source of the incompleteness could be traced directly to the expressibility in L of the predication operations. In this case the picture of the stages of incompleteness in first-order logic given in §25 would have to be complicated accordingly.
12. See Donald Davidson, 'Truth and Meaning'.
13. This difficulty has been pointed out in several places (e.g., Bruce Vermazen, 'Semantics and Semantics'). Efforts to repair the program have been made, but related difficulties also seem to beset them. For further discussions of this controversy, see, e.g., Davidson's 'Radical Interpretation' and papers in Evans and McDowell (eds.), *Truth and Meaning*.
14. The Quinean critics can, I believe, be answered in the fashion sketched in §5.
15. What I call Russell's theory is a synthesis of positions Russell took in the writings of his early period.
16. Frege's terms are 'Sinn' and 'Bedeutung'. On Feigl's translation they are rendered 'sense' and 'nominatum'; on Black's translation, 'sense' and 'reference'. I should mention, that, if 'Sinn' and 'Bedeutung' are read as new technical terms, perhaps Frege's theory would offend our semantic common sense somewhat less than it does when they are read as 'sense' and 'nominatum' (or 'sense' and 'reference').
17. I am not considering here the "coloring", conventional associations, or "connotation", in the literary sense, that these expressions might have.
18. Treating '*a*' or '*b*' as an incomplete symbol is the essence of the second part of Russell's answer to Frege's question. Russell's contextual definition of definite descriptions (and extensional abstracts) is really incidental. For one could eliminate instances of Frege's puzzle simply by treating definite-description and extensional-abstraction operators rather like quantifiers, i.e., as primitive *formula-producing* operators (versus singular-term producing operators).
19. For example, suppose that '$(\iota u)Au = (\iota x)Bx$' and '$(\iota u)Au = (\iota u)Au$' are true but different in meaning. Given a Russellian analysis, these sentences can be analysed as follows:

$$(\exists v)((Au \equiv_u v = u) \mathbin{\&} (\exists y)((Bx \equiv_x y = x) \mathbin{\&} v = y))$$
$$(\exists v)((Au \equiv_u v = u) \mathbin{\&} (\exists y)((Ax \equiv_x y = x) \mathbin{\&} v = y)).$$

In the result there are no expressions that have the same nominata but different meanings. True, the open sentences A and B are equivalent. But equivalence shows nothing since open sentences do not name anything; they only express.
20. I am optimistic about the viability of a Gricean defense of Russell. However, should this line of defense prove unsuccessful, that still would not spell defeat for what I am calling Russellian semantics. Even if definite descriptions were taken as semantically complete symbols, one would intuitively not want to say that they *name* anything. (Only names name.) They would be more like predicates and sentences: they would express something, and what they mean would be what they express. But what about referring? True enough, if definite descrip-

tions were semantically complete symbols, referring would seem to be a semantic relation. But it would be only a *derived* relation, defined as follows: θ refers to x *iff*$_{\mathrm{df}}$ (1) if θ is a definite description, then x = whatever unique object falls under what θ expresses, and (2) if θ is a name, then x = whatever θ names. This would be all there is to the commonsense theory of reference since predicates and sentences intuitively do not refer. On this account, then, there is still only one fundamental kind of meaning, and it partitions into naming and expressing. This is the essence of what I call Russellian semantics. Frege's '$a = a$'/'$a = b$' question could be answered within this framework. But our model structures would need to be enriched with a primitive definite description operation and appropriate predication and relativized predication operations.

21. See note 14, 'The Need for Abstract Entities in Semantic Analysis' and note 13, 'A Formulation of the Logic of Sense and Denotation'.
22. Intensional abstracts in L_ω with a Russellian semantics cannot give rise to instances of Frege's puzzle since $[A]_\alpha = [B]_\beta$, if true, is synonymous to $[A]_\alpha = [A]_\alpha$. Frege's semantics has an analogue of this feature. For example, let A_{o} and B_{o} be sentences; then, on the intended interpretation of the Frege-Church language $A_{o_1} = B_{o_1}$, if true, is synonymous to $A_{o_1} = A_{o_1}$.
23. Functional constants may be contextually defined in the usual way: $c_i^j(v_1, \ldots, v_j) =_{\mathrm{df}} (\iota v_{j+1})C_i^{j+1}(v_1, \ldots, v_j, v_{j+1})$. Here C_i^{j+1} is a predicate constant that expresses the $j+1$-ary relation-in-intension thought by Fregeans to be the Fregean sense of c_i^j. Conventions governing scope and the introduction of the new variable v_{j+1} are in force.
24. This does not rule out the possibility that co-denoting ordinary names 'a' and 'b'—and, in turn, materially equivalent sentences '$a = a$' and '$a = b$'—might have different coloring, stereotypical association, or connotation in the literary sense. Such differences, however, would not be differences in Fregean sense since the coloring, etc. of a name does not influence the determination of its nominatum. In a comprehensive theory of meaning it might be desirable to supplement the Russellian semantics with a second tier that deals with coloring, etc. Adding such a tier to a Fregean semantics would, by comparison, be much more difficult.
25. The use of $\mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}([A]_{v_1 \ldots v_j})$ in the definition is only shorthand; this is a defined notion whose definition is given in terms of $\mathscr{I}$, $\mathscr{A}$, $\mathscr{M}$ plus purely syntactic notions. The definition of $\mathrm{M}_{\mathscr{I}\mathscr{M}}$ can be given from scratch without appeal to $\mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}([A]_{v_1 \ldots v_j})$. The Russellian semantics for the language L from §37 is done analogously. Of course, certain further conditions (related to those set forth in §37) must be met by $\mathscr{I}$ and $\mathscr{M}$ in order to secure the specific interpretation intended for L. If, contrary to the thesis of extensionality, one were to take L_ω rather than L as canonical, then one would add to the Russellian semantics $\mathrm{M}_{\mathscr{I}\mathscr{M}}([A]_\alpha) = \mathrm{D}_{\mathscr{I}\mathscr{A}\mathscr{M}}([A]_\alpha)$, and one would add to the subsequent derivation of a Fregean semantics $\mathrm{R}_{\mathscr{I}\mathscr{M}}([A]_\alpha) = \mathrm{M}_{\mathscr{I}\mathscr{M}}([A]_\alpha)$ and $\mathrm{S}_{\mathscr{I}\mathscr{M}}([A]_\alpha) = \mathrm{M}_{\mathscr{I}\mathscr{M}}([v = [A]_\alpha]_v)$. (Here $[A]_\alpha$ is a closed term; if one wished to treat open terms, one would bring in the apparatus of assignments.)

In the setting of a ZF or GB-style theory for the predication relation, semantical paradoxes would rule out the definability (within the theory) of a Russellian meaning function for the theory itself. But one can give (within the theory) an implicit characterization by adjoining $M(\ulcorner A \urcorner) = [A]_{v_1 \ldots v_j}$ as an axiom schema, where A is any formula perhaps containing M itself. For more on this possibility, see §26.
26. Thus, "sense" determines "reference" via the "mode-of-presentation" function $\mathscr{G}$. Notice also that a theory of truth can easily be derived:

$$\mathrm{T}_{\mathscr{I}\mathscr{M}}(A) \textit{ iff}_{\mathrm{df}}\ \mathrm{R}_{\mathscr{I}\mathscr{M}}(A) = \mathrm{T} \textit{ iff}_{\mathrm{df}}\ \mathscr{G}(\mathrm{M}_{\mathscr{I}\mathscr{M}}(A)) = \mathrm{T}.$$

Incidentally, the Fregean theory of reference for predicates and open sentences is seen to be especially redundant when it is interpreted by means of the no-class applied set theory. (See §31.) For in that case, $R_{\mathscr{I}\mathscr{M}}(\theta) = \mathscr{G}(M_{\mathscr{I}\mathscr{M}}(\theta))$ says, in effect, only that θ "refers" to some intensional entity that is equivalent to the meaning of θ.

27. H. P. Grice, 'Meaning', 'Utterer's Meaning, Sentence Meaning, and Word Meaning', 'Utterer's Meaning and Intention', and Stephen Schiffer, *Meaning*. See also David Lewis, *Convention*, and Jonathan Bennett, *Linguistic Behavior*.
28. P. 215, Mates, 'Synonymity' (p. 125 in Linsky). I include under the heading 'Mates' puzzle', not just the particular puzzle given by Mates, but all analogous substitutivity puzzles involving synonymous predicates and formulas. See §18, where Mates' puzzle is distinguished from the paradox of analysis.
29. Geach, 'Intentional Identity'.
30. Kripke, 'A Puzzle About Belief'; John Perry, 'Frege on Demonstratives' and 'The Problem of the Essential Indexical'; David Kaplan, 'On the Logic of Demonstratives'; Tyler Burge, 'Belief and Synonymy' and 'Individualism and the Mental'.
31. Of course, these sentences might differ in coloring, literary connotation, etc.
32. Although this point might seem minor, it is important for metaphysics and epistemology. Perhaps there is a linguistic argument based solely on Mill's theory that wins the following conclusion:

If Hesperus = Phosphorus, then it is necessary that Hesperus = Phosphorus.

However, since 'H_2O' is not a name, such an argument could not be extended to show:

If water = H_2O, then it is necessary that water = H_2O.

To show this strong essentialist thesis, one must produce a different kind of argument, one that appeals to metaphysical, as well as linguistic, premises (as in the "twin-earth" argument in Putnam's 'The Meaning of "Meaning"'). What one must show to be necessary is not the trivial singular identity $\mathrm{Pred}_0(\mathrm{Pred}_0(\mathrm{Id}, H_2O), \mathrm{water})$—i.e., the proposition that water = water. We already knew that from Leibniz's law. Rather, one must show to be necessary a certain descriptive proposition concerning a complex relation among water, hydrogen, and oxygen. It is necessities of this latter kind that would require reforms in metaphysics. Mill's theory on its own does not support the existence of such necessities.

The descriptive character of expressions like 'H_2O', 'mean kinetic energy', etc. is also important for epistemology. There are conversational contexts in which it would be appropriate to utter both of the following sentences even though pot = marijuana:

(i) x knows that pot = pot.
(ii) x does not know that pot = marijuana.

The same holds for utterances of

(iii) x knows that water = water.
(iv) x does not know that water = H_2O.

even though water = H_2O. If Mill's theory of names is true, the ignorance reported in an utterance of (ii) would not originate in ignorance of relations among timeless descriptive concepts. Most likely it would originate in ignorance of a linguistic (or social or historical) matter. For example, x might smoke pot but not know what people speak of when they use the word 'marijuana' as a result of not knowing that 'marijuana' means pot. (I will not try to say here exactly what is involved in this kind of ignorance.) Therefore, if 'H_2O' were an ordinary name, then given Mill's theory the kind of ignorance reported in (iv) would also seem to be linguistic (social, historical) in origin. This epistemological conclusion would be striking, but it is not warranted since 'H_2O' is only a kind of definite description, not a name. When one discovers that water = H_2O (narrow scope), what one discovers is the correct chemical analysis of water. This type of discovery is essentially different from the type of discovery one makes when one discovers that pot = marijuana. Only the former non-linguistic (ahistorical, non-social) type of discovery constitutes basic progress in natural science, mathematics, and philosophy. Thus, Mill's theory on its own does not affect traditional epistemology in the ways some philosophers have thought.

33. As in Russell's 'Knowledge By Acquaintance and Knowledge By Description'.
34. Thus, 'conviction in acquaintance' is to be interpreted the way "internalists" interpret 'belief', and 'cognitive commitment', the way "externalists" interpret 'belief'. Both terms can be defined (using 'belief' and either 'epistemic access' or 'acquaintance'), but the definitions will depend on which interpretation of 'belief' is correct. The following schemas help further characterize the relation between cognitive commitments and convictions in acquaintance:

 x is cognitively committed to $[\ldots y \ldots]^y$ if x is convinced in acquaintance of $[\ldots y \ldots]^y$ or $[\ldots (\iota z)(Fz) \ldots]$, where $y = (\iota z)(Fz)$ and 'F' expresses one of x's modes of epistemic access to y.

 x is cognitively committed to $[\ldots G(\alpha) \ldots]$ if x is convinced in acquaintance of $[\ldots G(\alpha) \ldots]$ or $[\ldots \alpha\, \Delta\, (\iota z)(Fz) \ldots]$, where $\bar{G} = (\iota z)(Fz)$ and 'F' expresses one of x's modes of epistemic access to $\bar{G}$.

 Examples of what I mean by a mode of epistemic access include: meaning chains, social information chains, perceptual links, pictures, memory routes.
35. Just as a person can be acquainted with himself, a person can be acquainted with the present as the present is occurring. So suppose that I lose track of what time it is, that it is now noon, and that I erroneously believe that it is not now noon. Then, I am cognitively committed to $[u \text{ is not } u]^u$, where u is now (i.e., noon), and I have this necessarily false commitment in virtue of my not irrational conviction$_{acq}$ concerning, say, $[u$ is not mid-day on the day of which u is a moment$]^u$, where again u is now.

 In the perception example described earlier (the optical-apparatus example) our perceiver x is cognitively committed to $[y \neq y]^y$ in virtue of his conviction$_{acq}$ concerning, say, $[(\iota w)(x \text{ looks directly at } w) \neq (\iota w)(x \text{ looks at } w \text{ through a lens})]^x$. Again there is nothing irrational in this.
36. Tyler Burge expresses a closely related worry; see pp. 127 ff., 'Belief and Synonymy'; p. 97, 'Individualism and the Mental', and 'On Knowledge and Convention'. The issue here dramatizes the fact that any adequate theory of language learning must incorporate a resolution of the paradox of analysis.
37. In the amnesia example, the relevant conviction in acquaintance might be represented with something like the following: $[x \neq$ "<u>George Bealer" as he is called</u>$]^x$, where again I am x. Recall that metalinguistic propositions by no means exhaust the convictions in acquaintance that are pertinent to the puzzles about belief I am considering. Propositions concerning several other modes of epistemic access—e.g., historical information chains, perceptual links, memory routes, pictures, etc.—are also pertinent. For example, in Pierre's case one might want to represent his convictions$_{acq}$ as follows: [the city whose picture x remembers seeing is pretty]^x and [the city in which x resides is not pretty]^x, where x = Pierre.

38. This is the theory that would be associated with traditional Cartesianism.
39. For example, in order to handle certain modal puzzles involving belief, a proponent of the first position might want to strengthen the modal character of various candidate convictions$_{acq}$. E.g., in the amnesia example, someone might want to hold that $[x \neq$ "George Bealer" whoever he *in fact* is$]^x$, where I am x, is one of my convictions$_{acq}$. Given the special modal properties of the phrase 'whoever he *in fact* is', this proposition is necessary, not contingent. The proponent of the second position might not need to make this maneuver, for modals in natural language seem to attach to the asserted cognitive commitment rather than to the underlying conviction in acquaintance.
40. One can convert the suggested pragmatic solution into a formalized semantic solution. But to do so would be very artificial, for it would confuse the proper relation between semantics and conversational pragmatics. In any event, no theory of intentional language can succeed without invoking the acquaintance/description distinction at some stage or other. For it is needed to systematize an array of intuitive distinctions that are made in intentional language.
41. Incidentally, this anti-realist strategy seems to collide with a realist view advanced in 'Naming and Necessity'. Kripke tells us,

    > ...I hold the metaphysical view that, granted that there is no Sherlock Holmes, one cannot say of any possible person that he *would have been* Sherlock Holmes, had he existed. Several distinct possible people, and even actual ones such as Darwin or Jack the Ripper, might have performed the exploits of Holmes, but there is none of whom we can say that he would have *been* Holmes had he performed these exploits. For if so, which one? (p. 764)

    Now whenever people have beliefs about the same object, surely they could name it. Therefore, given Kripke's view that non-actual objects cannot be named, it follows that people cannot have beliefs about the same non-actual objects and, hence, that the anti-realist strategy is mistaken.

Chapter 8

1. Nothing in this section rides on whether these examples are right. Their purpose is only to impart the informal notions of quality and connection. In fact, I do not make formal use of the thesis that qualities and connections are special until the last chapter. None of the analyses given prior to that depend on this thesis.
2. Sidney Shoemaker has used the term 'Cambridge property' in this way (see his 'Causality and Properties'). This term derives from Geach's term 'Cambridge change' (see his *Logic Matters*, pp. 321 ff.; also *God and the Soul*, pp. 71 ff.).
3. My procedure here might strike one as circular since I am imparting the notion of a quality (connection) by appeal to the notion of a theoretical explanation, but good theoretical explanations are antecedently required to contain only vocabulary for genuine qualities (connections). This circle would be vicious if the aim of the procedure were to define what a quality (connection) is. But the aim here is not definition; it is only to introduce informally the notion of a quality (connection). Later in this chapter quality (connection) will be defined, not in terms of dependent notions like explanation, but in the setting of a purely logical theory.
4. Concerning the role of qualities in the explanation of change, many of Sidney Shoemaker's remarks in 'Causality and Properties' are congenial to my view. But I hold that not all qualities and connections are causal. Some are purely logical. And conceivably others are purely epiphenomenal (though I wish to remain neutral here on the existence of the latter).

5. I believe that it would not be appropriate here to engage in an extended consideration of the Wittgensteinian view that what makes for the objective existence of a quality (connection) is the common reaction of the members of a community to instances of it. Suffice it to say that the very notion of a "common reaction" presupposes the notion of a quality (connection); what makes two reactions to something truly similar is that they share qualities (connections). This holds for verbal reactions every bit as much as it does for non-verbal ones.
6. Pp. 72 ff., *Fact, Fiction, and Forecast*. Another role played by the concepts of quality and connection is in the statement of supervenience principles. And these concepts may also be expected to play a role in the analysis of the concept of randomness. Roughly, things that share a genuine quality form the ideal of a non-random aggregate; the random is that which grades off from this ideal.
7. This is not to say that there are not *special* circumstances in which some Cambridge properties and relations are projectible. The claim is that they are not projectible *ceteris paribus*. In any event, a complete explanation of why Cambridge properties and relations are projectible in some special circumstances (but not *ceteris paribus*) certainly requires an appeal to the underlying qualities and connections.
8. Someone might instead try to define qualities (connections) as those properties (relations) that support counterfactuals. Beyond various technical troubles facing this proposal, there is a methodological problem: it gets the proper order of explanation reversed. We ought to have some explanation of what makes some counterfactuals true and others false. Now, this can be done, but only by appealing to the concepts of quality and connection. Therefore, if these concepts are themselves explained in terms of their behavior in counterfactuals, we will be going in a circle. (Note the similarity to the circularity in the possible-worlds definition of necessity; see §46.) The only way out of the circle is to define the concepts of quality and connection within logic, specifically, within the logic for PRPs. According to this definition, what is logically distinctive about qualities and connections is their special role in the constitution of propositions.
9. Concerning philosophically basic forms, Plato was no doubt right that the dialectician can on his own "cut reality at its joints". But it seems that the dialectician is able to determine the true forms in *nature* only with assistance from the empirical scientist. How this might be done is the epistemological question I try to answer in the text.
10. Naturally, subjects of singular predications can belong to other metaphysical categories, e.g., to the categories of particular, matter, etc. (For example, the condition that figure (2) curves is a singular predication whose subject is figure (2), which is a particular.) Even though I must take account of this fact in my formal theory, I need not dwell on it in the present informal introduction.
11. The three relations *applying to*, *qualifying*, and *connecting* are all included in the converse of the Δ-relation.
12. Someone might wish to hold that thoughts and concepts fall under one of the intermediate conceptions between conception 1 and conception 2. Doing so would not require any alteration in the analyses given in the remainder of the book.
13. I am strongly attracted to the view that there are further metaphysical categories of primary predicables (such as actions, affections, places, times, stuffs) which play a role comparable to that of qualities and connections. If there are, it would not be difficult to adjust my theory to accommodate them.
14. For a word on Frege's representationalism see the sketch of his theory of meaning early in §38.
15. The language-of-thought theory is the most recent instance of representationalism. Of course, one could adhere to a certain version of this theory without being a representationalist, for one could just deny that anyone's thoughts ever

have any content at all. But the resulting theory would seem to be without content itself since *it* would presumably be thought by its proponents. Thus, unless some way were found to avoid this difficulty, this radical version of the theory would seem to be logically self-defeating. See note 7, chapter 1.

16. In order to get a better grasp of type 3 model structures, the reader might consider an artificial example: the qualities, connections, and conditions in $\mathscr{D}$ are genuine qualities, connections, and conditions, and the thoughts and complex concepts in $\mathscr{D}$ are identified with abstract trees (or sequences or abstract concatenations) whose ultimate constituents are these genuine qualities, connections, and conditions plus any of the condition-building operations. The range of the the thought-building operations would then be made up of these abstract trees (sequences, concatenations). Though such model structures might be helpful heuristically, there is perhaps reason to think that none of them yields a *natural* model. For what relations would serve to determine the *order* of the constituents in these abstract trees (sequences, concatenations)? If such ordering relations are not full-fledged connections, they must be Cambridge relations and, hence, complex concepts. But in this case one would have a regress on his hands. And technical feasibility aside, is it credible that thoughts are really trees, sequences, or concatenations?
17. So $|A|_{v_{i_1}}$ is a property abstract; the property it denotes is either a quality or a Cambridge property, depending on whether or not the property is necessarily equivalent to a quality. And $|A|_{v_1 \ldots v_m}$ is a relation abstract; the relation it denotes is either a connection or a Cambridge relation, depending on whether or not the relation is necessarily equivalent to a connection.
18. A thought-building operation f and a condition-building operation g are equivalent in the sense that $(\forall \alpha \in \mathscr{D})\mathscr{G}(f(\alpha)) = \mathscr{G}(g(\alpha))$. (In $\mathscr{L}$ with Δ, $(\forall \alpha)(\alpha \, \Delta \, f(\alpha) \equiv \alpha \, \Delta \, g(\alpha))$.)
19. This approach to the definition of correspondence would succeed even if thoughts and complex concepts fell under, not conception 2, but some intermediate conception between conceptions 1 and 2. The same thing can be said *mutatis mutandis* for the definitions of truth, necessity, analyticity, and validity given in the next chapter.

Chapter 9

1. P. 169, G. E. L. Owen, 'Logic and Metaphysics in Some Early Works of Aristotle'.
2. The same thing goes for the secondary uses of 'true' that arise in connection with speech acts and other intentional acts. I should mention that there are still other uses of 'true'—true north, a door true in its frame, being true to one's love, true gold, etc. I am willing to entertain the thought that the concepts associated with these uses—along with the concept of a true proposition itself—are all secondary concepts and that there is a single underlying primary concept of *fidelity to a thing's proper object*. This, however, would not vitiate the theory I am about to propose, for, in the case of a thought, fidelity to its proper object is just correspondence to a condition that obtains. However, in this event the various semantic concepts of truth would be no longer secondary; they would be tertiary.
3. This definition is adequate since, for every condition x, there is always a property y that has the required features. For example, let y be the property of

being identical to x if and only if x obtains (i.e., $y = |u = x \equiv x \text{ obtains}|^{x}_{u}$). Necessarily, this property has an instance (i.e., x) if and only if x obtains. And the condition that this property has an instance is necessarily equivalent to—and, hence, identical to—x itself.

4. Since a state of affairs is a condition that obtains, this definition is equivalent to: a thought is true iff_{df} it corresponds to a state of affairs.

   Incidentally, I have identified conditions that do not obtain with conception 1 entities, ontologically on a par with conditions that do obtain (i.e., states of affairs). However, just as Cambridge properties and relations are identical to mere complex concepts, conditions that do not obtain might be identical to mere thoughts. If so, the presentation in the text could be easily adjusted accordingly.

5. For example, there is a Zermelo-style theory in which the following truth schema is provable: $[A_u]$ is true $\equiv A_u$. I permit A_u to contain the definitions of all the various logical concepts I have been discussing: quality, connection, condition, thought, concept, correspondence, obtaining, truth, etc. No restrictions are imposed on occurrences of variables bound by quantifiers occurring within these definitions. However, I do require that all other occurrences of quantified variables in A_u are *grounded*; i.e., I require that these occurrences, which are the kind that might lead to paradoxes, be restricted in their range to the implicit universe of discourse u which is deemed relevant in the context. The above truth schema serves to justify the proposed definition of truth.

6. w is a possible world iff_{df} w is a world & $\neg(\exists y)(N[y \text{ is false}]^{y}$ & y is-true-in w). By the way, an analogous circularity would arise if the possible-worlds method were employed in a semantics for natural language, for in that case the idiom of possible worlds would again be used with realistic, not heuristic, intent.

7. I.e., x is just the condition that x is what it is.

8. $\mathscr{K}$ plays no role in type 2 algebraic model structures and, therefore, may be eliminated without loss there.

9. To be persuaded that this procedure is legitimate, consider the analogy with Tarski's uses of set theory as a background theory for his style of semantics. Incidentally, in the setting of a ZF or GB-style intensional logic, the Russellian meaning function for the theory itself cannot be defined outright within the theory, but it can be given an implicit characterization. See note 25, chapter 7.

10. Though in §14 I define validity for the language L_ω, my algebraic method works for a wide variety of languages L.

11. In 'Speaking of Objects' and 'Ontological Relativity' Quine maintains that the knowledge of the correct Tarskian reference function for a language is subject to rather the same sort of epistemological difficulties as the ones allegedly besetting knowledge of the correct meaning function. So he is not free to use epistemological grounds to reject the theory of meaning if, at the same time, he wants to accept the theory of reference.

12. By the same token, in view of Hume's remarks on relations of ideas, we may conclude that Hume rather approved of analyticity or something very close to it. However, if the Carnap-Lewis analysis were correct, that would make Hume's simultaneous approval of analyticity and attack on necessity (as well as Kant's defense of synthetic necessities) look pointless.

13. The general issue of whether the intentional act or its propositional object is primary will be discussed in the next section.

14. Again, I stress that someone could treat thoughts (and concepts) as falling under any of the intermediate conceptions between conceptions 1 and 2; analyticity could still be defined just as it is in the text.

    By the way, when Kant speaks of a concept as being "covertly" contained in another, I believe that he has in mind the sort of thing that shows up in instances of the paradox of analysis. If so, the apparatus used for resolving that paradox would no doubt need to be incorporated into a definition of analyticity that is fully faithful to Kant.

15. That is, propositions built up by means of one or more applications of the thought-building operations—conjunction, negation, ...
16. True, if Kant were to admit further categories of analytic thoughts, he would then be forced to countenance a host of analytic thoughts that are not self-evident (to us, at least). However, given that there are elementary instances of the paradox of analysis (see the previous note), it looks as though Kant could not avoid non-self-evident analyticities anyway. Furthermore, if a given analytic thought fails to be self-evident to a person at a time, it does not follow that this thought cannot become self-evident to the person later on. (Consider, e.g., the barber paradox.) Surely Kant would not say that the thought was once synthetic and that it subsequently became analytic.

    There is of course a full spectrum of "analyticity" concepts definable within the framework of the theory of qualities and concepts. But I know of no good reason for latching onto any intermediate concept in this spectrum.
17. Identity is defined in note 19 of chapter 4, and necessary equivalence is defined in §46.
18. Let 'F' express the Russell property $|x \not\Delta x|_x$. Does the fact that $[(\forall x)(\text{Lives}(x) \equiv x \,\Delta\, |\text{Lives}(x)|_x)]$ is valid imply that the paradoxical thought $[(\forall x)(Fx \equiv x \,\Delta\, |Fx|_x)]$ is valid? Not at all, for these two thoughts have different logical forms. The Russell property $|Fx|_x$ is a purely logical entity; $|\text{Lives}(x)|_x$ presumably is not. Further, $|Fx|_x$ is a Cambridge property and, hence, is logically complex; $|\text{Lives}(x)|_x$ is not a Cambridge property and, hence, is logically simple.
19. Equally serious obstacles arise because sentences are identified as the primary bearers of validity and because the Tarskian account must be given anew each time the primitive logical vocabulary of the language under consideration is enriched, even slightly.

    A general account of validity is also out of the question in ramified type theories such as Russell's or Church's. The concept of validity gives way to an infinite hierarchy of validity concepts. And the ramified theory is in principle unable to say what it is about them that makes them all validity concepts. This deficiency in ramified theories is much the same as the deficiency they encounter on the matter of truth. (See the closing remarks in §45.)
20. Occurrences of the thought-building and the condition-building operations might produce analogous obstacles to a Tarskian account of validity.
21. On this definitional strategy I begin with a stipulated list of fundamental logical operations. There is, however, an alternate strategy which deserves mention. If one took the concepts of thought, concept, quality, connection, and condition as primitive, then one could define a logical operation to be a connection that can be used to build in an appropriate way thoughts, concepts, qualities, connections, or conditions from other thoughts, concepts, etc.
22. There are any number of restricted "validity" concepts that could be defined in the theory of qualities and concepts. But there seems to be no good rationale for drawing the line at one of them. For this reason, I see no alternative but to accept the unrestricted concept defined in the text.

    A possible objection to this definition is that according to it many valid thoughts would seem to lack the epistemic properties—self-evidence, aprioricity—traditionally attributed to valid thoughts. However, what is self-evident to a person and what he knows *a priori* depend on his intelligence and intellectual training. Yet one's intelligence can grow, and one can educate one's logical intuitions. So what is *actually* self-evident at a given time should not be used as a test of what counts as valid. In addition if *possible* self-evidence were taken as a test, it is unclear that it would count against my definition of validity. For it is unclear that any validities on my conception necessarily fail to be self-evident or knowable *a priori* for everyone.

23. Church's theorem also helped to transform the Aristotelian conception of logic.

Chapter 10

1. 'Intentionality', vol. iv, *The Encyclopedia of Philosophy*.
2. Recall that complex ideas are intensional entities that can be analysed by means of the inverses of the thought-building operations; simple ideas are the intensional entities that cannot.
3. The concept of logical form is characterized in §47.
4. This is the point of contact between the general phenomenon of intentionality and Geach's special phenomenon of intentional identity. Suppose that in actual conversation I say that a certain complex idea is about a particular object that I know does not exist. Then, according to the descriptivist analysis suggested near the end of §39, the assertion does not entail that there is an intentional inexistent, non-existent subsistent, or non-actual possible such that the idea is the result of predicating a concept of this supposed object; rather, it only entails that the idea has a certain pragmatically determined descriptive character such that the associated descriptive concept does not apply to anything. It is important to realize, however, that the analysis I give of intentionality in no way requires that there are no non-actual objects. Their existence is an independent issue.
5. Alternatively, if M_2 were synonymous to the intentional prefix 'It is necessary that John believes that', that would also show that the definition does not provide a necessary condition.
6. Note that the sentences M_5(quarks have charm) and M_5(the physical line has a well-ordering) would also seem not to be contingent. Furthermore, this would be so even if M_5 were instead synonymous to the intentional prefix 'John perceives with his senses that'.
7. For simplicity of exposition, I will proceed under the assumption that factive intentional relations are not connections. If there are factive intentional connections—and there is some reason to think there are—then step one in the text should be divided into two parts: first, the concept of a non-factive intentional connection is defined; then, in terms of this narrower concept the general concept of an intentional connection—factive or non-factive—is defined. For example, a connection is intentional *iff* it is necessarily included in a non-factive intentional connection. I might also state explicitly here that intentional connections are never necessarily null.
8. I suspect that Chisholm had something rather like this in the back of his mind when he attempted to define what it takes for a *simple* sentence prefix to be intentional.
9. In symbols:

$$
\begin{array}{l}
\text{a connection } x \text{ is intentional } \mathit{iff}_{\mathrm{df}} \\
\Diamond(\exists y, z)(\text{Individual}(y)\ \&\ \text{Complex Idea}(z)\ \& \\
\Diamond(\langle y, z\rangle\ \Delta\ x\ \&\ (\text{True}(z) \vee (\exists w) w\ \Delta\ z))\ \& \\
\Diamond(\langle y, z\rangle\ \Delta\ x\ \&\ \neg(\text{True}(z) \vee (\exists w) w\ \Delta\ z))\ \& \\
\Diamond\langle y, z\rangle\ \not\Delta\ x).
\end{array}
$$

Notice that the analysis in no way restricts the range of the connection x. As far as the analysis is concerned, items from any metaphysical category can be in the range of x. Notice also that the analysis is given for binary connections only; however, it can be generalized. Finally, recall that in the text I am proceeding under the simplifying assumption that intentional connections are never factive. If there are factive intentional connections, then the definition in the text should be understood as a definition of the narrower concept of a non-factive intentional connection, and the general concept of an intentional connection, factive or non-factive, would then be defined in terms of this concept (as, e.g., in note 7).

10. The same consideration would rule out as a counterexample the empirical relation holding between x and n such that x is n grams (in mass). Furthermore, this relation would fail as a counterexample for the additional reason that it is a Cambridge relation, not a connection.

11. It would be a backward strategy to try repairing the counterexample by allowing complex ideas to be in the relation's range, e.g., to propose as a counterexample the dispositional relation $|y$ is a thought and x is a hunk of salt having the disposition to dissolve whenever y is true$|_{xy}$. This is certainly not a counterexample, for it is no connection, causal or otherwise. Hunks of salt are just not connected causally—or in any other modally relevant way—to thoughts. Only by thinking can an individual particular bear a modally relevant connection to thoughts.

    The relation holding between a particular and its Aristotelian final cause (e.g., between a kidney and the process of filtering wastes) might seem to be an exception to what I have just said. I would disqualify this relation as a counterexample to my analysis of intentionality on two counts. First, final causes are, it seems, not complex ideas. Secondly, the relation does not meet the contingency requirement in the analysis, for it holds necessarily, not contingently, between a particular and its Aristotelian final cause (if the Aristotelian theory is correct). For example, if the Aristotelian final cause of a kidney is filtering wastes, then it could not fail to have this as its Aristotelian final cause. There is, of course, a contingent relation that holds between particulars and their accidental uses or functions. (E.g., some hammers are, as a contingent fact, used for holding down papers.) But this relation relates an individual and a use only if that use actually has an instance. (E.g., necessarily, a hammer is used for holding down papers only if papers are sometimes actually held down by it.) Thus this relation does not meet the independent-veracity requirement in the analysis. There is, finally, a relation holding between particulars and their intended accidental uses. (E.g., I might intend to use a hammer for holding down papers even though I have not yet done so.) But this relation is no counterexample since it is an intentional relation. (The token-meaning relation is a special case of this kind of intentional relation; see candidate counterexample (4).)

12. A further reason why utterance-token meaning fails to be a counterexample is that, like speaker meaning, it is an intentional relation. But if this relation were truly a connection, I should not be happy to leave things here. The goal is to get at those connections simply in virtue of which a creature (as opposed to an utterance token produced by a creature) is intentional, and utterance-token meaning would not be such a connection. Thus, if it were conceded that utterance-token meaning is a connection, I should want to implement the following routine. A plurality of intentional connections is deemed *minimal* if and only if all intentional connections can be defined in terms of them, where no smaller number of them suffices for this purpose. Then, an intentional connection is defined to be *psychological* if and only if it is necessarily included in one of the connections in such a minimal plurality. These psychological connections are

those in virtue of which a creature is intentional. Since utterance-token meaning is not psychological in this sense, this routine would solve the problem. I do not implement this routine in the text, for, after all, utterance-token meaning seems disqualified as a connection in the first place.

In speaking of utterance-token meaning, I have been referring to a meaning relation that holds contingently between utterance tokens and intensions in virtue of appropriate utterer's intentions. What is the relationship between intentional meaning relations and Russellian meaning functions $M_{\mathscr{J}}$ which arise in formal semantics? The answer is roughly this. In order to establish a given language as their own, the members of a speech community institute a convention to produce tokens of utterance types in the language only if the Russellian meaning (as specified by the $M_{\mathscr{J}}$ for the language) of the utterance type is the intension that the speaker (intentionally) means when he produces the token. Because a Russellian meaning function $M_{\mathscr{J}}$ is an antecedently given purely abstract pairing of utterance types and intensions, it does not qualify as a counterexample to the analysis of intentionality. First, $M_{\mathscr{J}}$ pairs utterance types (i.e., universals) and intensions, but to be a counterexample it would have to pair particulars and intensions. Secondly, $M_{\mathscr{J}}$ violates the contingency requirement since it is L-determinate. Thirdly, since $M_{\mathscr{J}}$ is an arbitrary pairing of utterance types and intensions, there is no reason to think that it is a genuine connection.

13. For example, the relation of coexisting—i.e., $[(\exists u)x = u \equiv (\exists u)y = u]_{xy}$—might be offered as a counterexample. I would deny that this relation is a true connection. But if someone is in doubt about this, he could simply tighten up the modalities in the analysis by requiring that the contingency alluded to there have some source other than the contingent existence of individual particulars. The resulting analysis would be immune to the candidate counterexample.
14. I will offer one analysis that I find plausible. There are promising alternatives, however. What matters to me is that there exist at least one purely logical analysis of the mental that is possible along the suggested or related lines.
15. The following more complex analysis also suggests itself:

$$(\forall x)(uSx \supset (\exists y)(vSy \;\&\; x \text{ is similar to } y)) \;\&$$
$$(\forall x)(vSx \supset (\exists y)(uSy \;\&\; x \text{ is similar to } y))$$

where S expresses either the sensing relation or the relation holding between beings and their own sensory events. But in what way are items x and y supposed to be *similar*? Not in their physiological underpinnings; the sense experience of creatures u and v could in principle be qualitatively alike even though u and v are quite different physiologically. Not in their behavioral correlates; u and v could behave similarly even though their sense experience is qualitatively different, and u and v could behave differently even though their sense experience is qualitatively the same. And not functionally; sense experiences that differ qualitatively could have the same functional roles, and qualitatively similar sense experiences could in principle function differently. Evidently, the way in which the sense experiences of u and v must be similar is in the qualities of which u and v are sensorily aware, i.e., in the *sensible qualities* sensed by u and v. However, this is precisely the conclusion I am driving toward in the text. For simplicity I will therefore not pursue the more complex analysis suggested in this note.

16. This conclusion is not essential to my argument. I include it in the text as a foil for the point that sensible qualities and sensible conditions are public objects. The analysis I finally give is strictly speaking neutral with regard to the existence of private mental particulars.
17. The theory that qualities are the objects of the sensing relation is a theory

advocated by Russell; see p. 17 in *The Problems of Philosophy* (p. 12, paperback edition). I am allowing in addition that conditions might be sensed. I stress that this possibility is not essential to anything I say here or below and that all traces of it may be stricken from my official analysis.

18. Sense-data theorists, who would deny this conclusion, could base a purely logical analysis of sensation on the special ontological dependence that sense data are supposed to have upon sensible qualities (if x has a sense datum that is red, then that particular sense datum exists only as long as x senses red); the overall analysis of the mental would then be adjusted accordingly.
19. This fact about feeling seems to have an intentional analogue involving self-consciousness. For example, it seems that x can have the thought that x has the thought that A only if the thought that x has the thought that A is *true*. If so, inner feeling and self-consciousness are analogous in that they both *reflect* one's own current state. In this, they differ both from sensing, which is the lowest level of the psyche, and from non-self-conscious thinking, which is the lowest level of reason. Hence, a logical basis for the classical partitioning of the psyche.
20. If this alternative is incorrect, then inner feeling should be analysed directly using the logically distinctive feature just cited in the text.
21. In addition to reflective qualities there might be reflective connections and conditions. The reflective connections would be none other than the conscious mental connections, the conscious operations of mind. My final analysis of the concept of a mental connection is compatible with this view of reflective connections and conditions. At the same time, the analysis does not depend on it.
22. Recall that I am in the process of trying to say what pure experience is supposed to be. It need not be assumed here that there is such a thing as pure experience. Nor for that matter must it be assumed that there even could be such a thing. Therefore, in the text I should not be understood as asserting that there are these two types of basic mental phenomena, pure experience and thinking. Rather, I should be understood as analysing what one believes when one believes that there are.
23. Perhaps I should also make it explicit that mental connections are never necessarily null. Notice that the analysis has been phrased so as to cover factive intentional connections if there are any. (See notes 7 and 9.) And notice that the analysis imposes no restriction on the range of mental connections; as far as the analysis is concerned, items from any metaphysical category could be in the range of any mental connection. There are numerous adjustments that one could make in the analysis so as to meet certain worries that might arise. The thesis to which I am committed is that some variant of the analysis is free of all counterexamples. As before, my hope is that the reader will see how particular adjustments would be made. Finally, by using the analysis of what a mental connection is, one can go on to define what a Cambridge mental relation is. Putting the two definitions together, one will have succeeded in characterizing what a mental relation is in general.
24. Should the critic hesitate to accept this reply, there is a way to adjust the analysis so that it guarantees that no non-mental causally grounded dispositional relations satisfy it. Associated with each dispositional relation is a kernel non-dispositional relation. To see what I mean by this, consider the dispositional relation holding between particulars x and qualities y such that x is *disposed* to be achieved physically by instances of y. Associated with this dispositional relation is the kernel non-dispositional relation holding between particulars x and qualities y such that x is *actually* activated physically by instances of y. This kernel relation is unable to hold between a particular and a quality unless that

quality actually has an instance, so this relation does not satisfy the analysis of mental connection given in the text. What is distinctive about mental dispositional connections is that their associated non-dispositional kernels do satisfy the analysis given in the text. This difference between mental and non-mental dispositions could be made use of if the critic felt a need to adjust the analysis of the mental in order to avoid the threat of dispositional counterexamples.

25. Suppose $|\langle v_1, \ldots, v_n \rangle \,\Delta\, z|$ is a phenomenon, where $n \geq 1$ and z is a quality or connection different from the special logical qualities and connections characterized a moment ago. Then, in order for this phenomenon to be purely physical, z must be a physical quality or connection. Thus, a phenomenon would be purely physical only if it is not the phenomenon $|\langle x, y \rangle \,\Delta\, z|$, where x is a particular, y an idea, and z a mental connection.
26. I advocate treating the concepts of stuff, location, and time as purely logical concepts. See 'Predication and Matter' for a sketch of how one might go about doing this.
27. In 'An Inconsistency in Functionalism' I show that mental relations have adequate functional definitions if and only if they have adequate ordinary explicit definitions, behavioral or physiological. Since there are independent reasons for thinking there are counterexamples to the definitions provided by behaviorism and naive physiological reductionism, my result in effect predicts that there are counterexamples to the definitions provided by functionalism. Indeed, the predicted counterexamples seem to exist. For example, just as believing and the Cambridge relation pretending-to-believe (or wanting and the Cambridge relation pretending-to-want) might in an appropriate circumstance be behaviorally indistinguishable, it seems that they might at least in principle be functionally indistinguishable as well. And Ned Block's China example (p. 279 ff., 'Troubles With Functionalism') also seems to be a counterexample to functional definitions of mental relations. Now there has been difficulty in saying why the China example is really a counterexample. My analysis of the mental provides an answer: the functionally defined relations in the example are only Cambridge relations whereas the mental relations that the functionalist must capture in his functional definitions are genuine connections. (Connections are those relations that fix the causal, phenomenal, and logical order of the world; the China relations are not like this.) This observation suggests that functional definitions might be salvaged by restricting the values of their relation variables to genuine connections. But this move, while it might help functionalism, would only hurt materialism. For in the setting of any adequate psychological theory for intentional matters, this tightening of functional definitions would amount to restricting the values of the relation variables to connections that have just those logical features that are, on my analysis, distinctive of mental connections. In the same vein, the fact that functional definitions use only physical (and logical) terminology does not show that functionalism is a form of naturalism; all sorts of non-naturalistic things can be expressed using only physical (and logical) terminology. Physicalism (of the terminological kind) entails neither materialism nor naturalism.
28. It is not essential to assume here or in what follows that the particular relations just cited are in fact genuine connections.
29. In symbols, $|x \text{ is conscious of } y|_{xy} = |(\exists z)(z \text{ is a mental connection } \& \langle x, y \rangle \,\Delta\, z)|_{xy}$.
30. The question of whether pure experiencing (or sensing or feeling) should be classified with these conscious connections may be left open in the present discussion. However, I strongly doubt that it should be classified with them.

(Most functionalists today would concur.) The argument is analogous to the argument given earlier against "non-conscious consciousness".

31. It is not essential to assume here or in what follows that these relations are genuine connections or that they are distinct from one another. The present topic could strictly speaking be by-passed; however, it is helpful heuristically.

32. Conscious connections thus fall into two kinds—those that are, and those that are not, functionally significant when looked at mechanically as if no consciousness were involved. The former, but not the latter, can be paired with non-conscious functional analogues (if non-conscious mental connections truly exist). Since the consciousness relation is the mental connection that necessarily includes all conscious connections, consciousness may be defined as follows: consciousness $=_{\text{df}}$ relating by means of a mental connection which, relative to a pairing of functional analogues, necessarily includes one mental connection in each pair of paired mental connections and also all unpaired mental connections. More formally, $|x$ is conscious of $y|_{xy} =_{\text{df}} |(\exists z)(z$ is a mental connection & $(\exists t)(t$ is a 1-1 function & relations paired by t are functional analogues & $(\forall w)((w$ is a mental connection & either w is in the domain of t or w is not in the field of $t)$ $\supset \Box w \subseteq u))$ & $\langle x, y\rangle \Delta u)|_{xy}$. This definition can be rendered within the theory of qualities and concepts once the concept of a pairing t of functional analogues is suitably approximated within the theory. Something like the following condition on t should do: for all thoughts r and s, if the thought-building trees of r and s are just alike except that in r's tree the predicate nodes are relations in the domain of t and in s's tree the predicate nodes are the corresponding relations in the range of t, then r and s have the same modal value (necessary, contingent, impossible). The following much weaker condition on t also seems to suffice for the purpose of defining consciousness: for all relations w and w' paired by t, the potential range of w is the same as the potential range of w'. The definition of consciousness I give in the text is simpler than the above, however, and it has metaphysical and historical appeal of its own.

33. These category limitations do not all fall along the proposition/non-proposition line. E.g., the ranges of the non-factive mental connections (such as believing) differ from the ranges of any that might be factive (such as knowing). Additional category restrictions seem to apply within the non-factive relations themselves; e.g., although one can wish one were presently unconscious, it seems that one cannot really be convinced that one is presently unconscious. In a related vein, it seems that one cannot really decide (to bring it about) that green is red. (In any case, one certainly cannot decide upon truths in the same sense as one can decide upon courses of actions, nor can one in this sense decide upon oneself.) Similarly, one cannot really desire (to bring it about) that 1 is a number (though one can desire that it be true that 1 is a number). Finally, though there might be a sense in which one can want cheese, in no sense can one believe cheese.

34. In symbols, a mental relation is transcendental *iff*$_{\text{df}}$ $(\forall v)(\Diamond(\exists u, w)(w$ is a mental connection & $\langle u, v\rangle \Delta w) \supset \Diamond(\exists u)\langle u, v\rangle \Delta z)$.

35. To be conscious of the self and the present would not be to be conscious of special qualities (as is the case in sensation) if St. Augustine is right.

36. The metaphysical picture of mental relations sketched in §39 might call for certain adjustments in the definition. Notice also that the definition does not strictly state that consciousness is a genuine connection. It is, however, for it is what explains the unified awareness in our mental lives.

References

Aristotle, *The Works of Aristotle Translated into English*, vol. viii, ed. W. D. Ross (12 vols., London, 1908–1952).

——, *Aristotle's Categories and de Interpretatione*, trans. J. L. Ackrill (Clarendon Aristotle Series), Oxford, 1963.

Bar-Hillel, Y., ed., *Proceedings of the 1964 International Congress for Logic, Methodology and Philosophy of Science* (Amsterdam, 1965).

——, Poznanski, E., Rabin, M., and Robinson, A., eds., *Essays on the Foundations of Mathematics, Dedicated to A. A. Fraenkel on his Seventieth Anniversary* (Amsterdam, 1962).

Bealer, G., 'Predication and Matter', *Synthese*, xxxi (1975), 493–508. Reprinted in Pelletier.

——, 'An Inconsistency in Functionalism', *Synthese*, xxxviii (1978), 333–72.

——, 'Theories of Properties, Relations, and Propositions', *The Journal of Philosophy*, lxxvi (1979), 634–48.

——, 'Foundations Without Sets', *American Philosophical Quarterly*, xviii (1981), 347–53.

——, 'Completeness in the Theory of Properties, Relations, and Propositions', *The Journal of Symbolic Logic*, xlviii (1983), 415–26.

Benacerraf, P., 'What Numbers Could Not Be', *The Philosophical Review*, lxxiv (1965), 47–73.

——, 'Mathematical Truth', *The Journal of Philosophy*, lxx (1973), 661–79.

——, and Putnam, H., eds., *Philosophy of Mathematics, Selected Readings* (Englewood Cliffs, 1964).

Bennett, J., *Linguistic Behavior* (Cambridge, 1976).

Black, M., ed., *Philosophy in America* (Ithaca, 1965).

Block, N., 'Troubles With Functionalism', in Savage, 261–325.

Boolos, G., 'The Iterative Conception of Set', *The Journal of Philosophy*, lxviii (1971), 215–31.

Brentano, F., *Psychologie vom empirischen Standpunkt* (Vienna, 1874). Translated in part as 'The Distinction Between the Mental and Physical' in Chisholm, *Realism and the Background of Phenomenology*.

Burge, T., 'Knowledge and Convention', *The Philosophical Review*, lxxxiv (1975), 249–55.

——, 'Belief and Synonymy', *The Journal of Philosophy*, lxxv (1978), 119–38.

——, 'Semantical Paradox', *The Journal of Philosophy* lxxvi (1979), 169–98.

——, 'Individualism and the Mental', *Midwest Studies in Philosophy*, iv (1979), 73–122.

Carnap, R., *Der logische Aufbau der Welt* (Berlin, 1928). Translated with additions as *The Logical Structure of the World and Pseudoproblems in Philosophy* (Berkeley, 1967).

——, *Logische Syntax der Sprache* (Vienna, 1934). Translated with additions as *The Logical Syntax of Language* (London, 1937). Paperback (Paterson, 1959).

——, *Meaning and Necessity* (Chicago, 1947).

Chihara, C., *Ontology and the Vicious-Circle Principle* (Ithaca, 1973).

Chisholm, R., ed., *Realism and the Background of Phenomenology* (Glencoe, 1960).

——, 'Intentionality', in Edwards, iv, 201–4.

——, and Sellars, W., see Sellars.

Church, A., Review of Morton White's 'A Note on the "Paradox of Analysis"', Max Black's 'The "Paradox of Analysis" Again: a Reply', Morton White's 'Analysis and Identity: a Rejoinder', and Max Black's 'How Can Analysis Be Informative?', *The Journal of Symbolic Logic*, xi (1946), 132–3.

——, 'On Carnap's Analysis of Statements of Assertion and Belief', *Analysis*, x (1950), 97–9.

——, 'The Need for Abstract Entities in Semantic Analysis', *Proceedings of the American Academy of Arts and Sciences*, lxxx (1951), 100–12.

——, 'A Formulation of the Logic of Sense and Denotation', in Henle, Kallen, and Langer, 3–24.

——, 'Intensional Isomorphism and Identity of Belief', *Philosophical Studies*, v (1954), 65–73.

——, 'Outline of a Revised Formulation of the Logic of Sense and Denotation', in two parts, *Noûs*, vii (1973), 24–33, and viii (1974), 135–56.

——, 'Comparison of Russell's Resolution of the Semantical Antinomies with that of Tarski', *The Journal of Symbolic Logic*, xli (1976), 747–60.

Davidson, D., 'Theories of Meaning and Learnable Languages', in Bar-Hillel, *Proceedings of the 1964 International Congress for Logic, Methodology and Philosophy of Science*, 383–94.

——, 'Truth and Meaning', *Synthese*, xvii (1967), 304–23.

——, 'On Saying That', *Synthese*, xix (1968), 130–46.

——, 'Radical Interpretation', *Dialectica*, xxvii (1973), 313–28.

——, and Harman, G., eds., *Semantics of Natural Language* (Dordrecht, 1972).

——, and Harman, G., eds., *The Logic of Grammar* (Encino, 1975).

Dold, A., and Eckmann, B., eds., *Proof Theory Symposium, Kiel 1974* (Lecture Notes in Mathematics, vol. 500), Berlin, 1975.

Donnellan, K., 'Reference and Definite Descriptions', *The Philosophical Review*, lxxv (1966), 281–304.

Dummett, M., *Elements of Intuitionism* (Oxford, 1977).

Düring, I., and Owen, G. E. L., eds., *Aristotle and Plato in the Mid-Fourth Century* (Göteborg, 1960).

Edwards, P., ed., *The Encyclopedia of Philosophy* (8 vols., New York, 1967).

Enderton, H., *A Mathematical Introduction to Logic* (New York, 1972).

Evans, G., and McDowell, J., eds., *Truth and Meaning: Essays in Semantics* (Oxford, 1976).

Feferman, S., 'Non-Extensional Type-Free Theories of Partial Operations and Classifications', in Dold and Eckmann, 73–118.

Feigl, H., and Sellars, W., eds., *Readings in Philosophical Analysis* (New York, 1949).

——, Scriven, M., and Maxwell, G., eds., *Concepts, Theories, and the Mind-Body Problem* (Minnesota Studies in the Philosophy of Science, vol. iv), Minneapolis, 1958.

Fitch, F., 'The System CΔ of Combinatory Logic', *The Journal of Symbolic Logic*, xxviii (1963), 87–97.

——, 'A Consistent Combinatory Logic with an Inverse to Equality', *The Journal of Symbolic Logic*, xliv (1980), 529–43.

Frege, G., *Die Grundlagen der Arithmetik* (Breslau, 1884). Reprinted with translation as *The Foundations of Arithmetic* (New York, 1950).

——, *Funktion und Begriff* (Jena, 1891). Translated as 'Function and Concept' in Frege, *Translations from the Philosophical Writings of Gottlob Frege*.

——, 'Über Begriff und Gegenstand', *Vierteljahrschrift für wissenschaftliche Philosophie*, xvi (1892), 192–205. Translated as 'On Concept and Object' in Frege, *Philosophical Writings*.

——, 'Über Sinn und Bedeutung', *Zeitschrift für Philosophie und philosophische Kritik*, c (1892), 22–50. Translated as 'On Sense and Nominatum in Feigl and Sellars and as 'On Sense and Reference' in Frege, *Philosophical Writings*.

——, *Grundgesetze der Arithmetik* (2 vols., Jena, 1893, 1903). Translated in part as *The Basic Laws of Arithmetic* (Berkeley, 1967).

——, *Translations from the Philosophical Writings of Gottlob Frege*, eds. P. T. Geach and M. Black (Oxford, 1952).

French, P., Uehling, T., and Wettstein, H., eds., *Contemporary Perspectives in the Philosophy of Language* (Minneapolis, 1979).

Friedman, H., 'The Consistency of Classical Set Theory Relative to a Set Theory with Intuitionistic Logic', *The Journal of Symbolic Logic*, xxxviii (1973), 315–19.

Gallin, D., *Intensional and Higher-Order Modal Logic* (Amsterdam, 1975).

Gandy, R. O., 'On the Axiom of Extensionality', part II, *The Journal of Symbolic Logic*, xxiv (1959), 287–300.

Geach, P. T., 'Intentional Identity', *The Journal of Philosophy*, lxiv (1967), 627–32.

——, *God and the Soul* (New York, 1969).

——, *Logic Matters* (Oxford, 1972).

Gilmore, P. C., 'The Consistency of Partial Set Theory Without Extensionality', *Proceedings of Symposia in Pure Mathematics*, viii, part II (1974), 147–53.

Gödel, K., 'Russell's Mathematical Logic', in Schilpp, 125–53. Reprinted in Benacerraf and Putnam.
Goodman, N., *Fact, Fiction, and Forecast* (Cambridge, Mass., 1955).
Grandy, R., 'Anadic Logic and English', *Synthese*, xxxii (1976), 395–402.
Grice, H. P., 'Meaning', *The Philosophical Review*, lxvi (1957), 377–88.
——, 'Utterer's Meaning, Sentence Meaning, and Word Meaning', *Foundations of Language*, iii (1968), 225–42.
——, 'Utterer's Meaning and Intentions', *The Philosophical Review*, lxxviii (1969), 147–77.
——, 'Logic and Conversation' in Davidson and Harman, *The Logic of Grammar*, 64–74.
——, 'Definite Descriptions in Russell and in the Vernacular', unpublished.
Gunderson, K., *Language, Mind, and Knowledge* (Minnesota Studies in the Philosophy of Science, vol. vii), Minneapolis, 1975.
Halmos, P., *Naive Set Theory* (Princeton, 1960).
Hambourger, R., 'A Difficulty with the Frege-Russell Definition of Number', *The Journal of Philosophy*, lxxiv (1977), 409–14.
Henkin, L., Monk, D., and Tarski, A., *Cylindric Algebras, Part I* (Amsterdam, 1971).
Henle, P., Kallen, H. M., and Langer, S. K., eds., *Structure, Method, and Meaning: Essays in Honor of Henry M. Scheffer* (New York, 1951).
Hintikka, J., *Models for Modalities* (Dordrecht, 1969).
——, Moravcsik, J., and Suppes, P., eds., *Approaches to Natural Language: Proceedings of the 1970 Stanford Workshop on Grammar and Semantics* (Dordrecht, 1973).
Hume, David, *A Treatise of Human Nature*, ed. L. A. Selby-Bigge (Oxford, 1888).
Inwagen, P. van, ed., *Time and Cause, Essays Presented to Richard Taylor*, to be published.
Kant, Immanuel, *Critique of Pure Reason*, trans. N. K. Smith (New York, 1965).
Kaplan, D., 'Quantifying In', *Synthese*, xix (1968–9), 178–214.
——, 'What Is Russell's Theory of Descriptions?', in Yourgrau and Breck, 277–88.
——, 'How to Russell a Frege-Church', *The Journal of Philosophy*, lxxii (1975), 716–29.
——, 'On the Logic of Demonstratives', in French, Uehling, and Wettstein, 401–14.
Kelly, J., *General Topology* (New York, 1955).
Kessler, G., 'Frege, Mill, and the Foundations of Arithmetic', *The Journal of Philosophy*, lxxvii (1980), 65–79.
Klibansky, R., ed., *Contemporary Philosophy: A Survey* (Florence, 1968).
Kripke, S., 'Naming and Necessity', in Davidson and Harman, *Semantics*, 253–355 and 763–9.
——, 'Outline of a Theory of Truth', *The Journal of Philosophy*, lxxii (1975), 690–716.
——, 'A Puzzle About Belief', in Margalit, 239–83.

Lewis, D., *Convention* (Cambridge, Mass., 1969).
——, 'General Semantics', in Davidson and Harman, *Semantics* 169–218.
——, *Counterfactuals* (Cambridge, Mass., 1973).
Linsky, L., ed., *Semantics and the Philosophy of Language* (Urbana, 1952).
Locke, John, *An Essay Concerning Human Understanding*, vol. i, ed. A. C. Fraser (New York, 1959).
Marcus, R. B., 'Classes, Collections, and Individuals', *American Philosophical Quarterly*, xi (1974), 227–32.
Margalit, A., ed., *Meaning and Use, Papers Presented at the Second Jerusalem Philosophical Encounter* (Dordrecht, 1979).
Mates, B., 'Synonymity', *University of California Publications in Philosophy*, xxv (1950), 201–26. Reprinted in Linsky.
Montague, R., 'Pragmatics', in Klibansky, 95–118.
——, 'Universal Grammar', *Theoria*, xxxvi (1970), 373–98.
——, 'The Proper Treatment of Quantification in English', in Hintikka, Moravcsik, and Suppes, 221–42.
Myro, G., 'Aspects of Acceptability', *Pacific Philosophical Quarterly*, lxii (1981), 107–22.
Owen, G. E. L., 'Logic and Metaphysics in Some Early Works of Aristotle', in Düring and Owen, 163–90.
Parsons, C., 'Frege's Theory of Number', in Black, 180–203.
——, 'The Liar Paradox', *Journal of Philosophical Logic*, iii (1974), 381–412.
——, 'Intensional Logic in Extensional Language', *The Journal of Symbolic Logic*, xlvii (1982), 289–328.
Pelletier, F. J., ed., *Mass Terms* (Dordrecht, 1979).
Perry, J., 'Frege on Demonstratives', *The Philosophical Review*, lxxxvi (1977), 474–97.
——, 'The Problem of the Essential Indexical', *Noûs*, xiii (1979), 3–21.
Pitcher, G., ed., *Truth* (Englewood Cliffs, 1964).
Plato, *The Collected Dialogues of Plato*, eds. E. Hamilton and H. Cairns (Princeton, 1961).
Putnam, H., 'The Meaning of "Meaning"', in Gunderson, 131–91.
Quine, W. V. O., 'New Foundations for Mathematical Logic', *American Mathematical Monthly*, xliv (1937), 70–80. Reprinted with additions in Quine, *From a Logical Point of View*.
——, *Mathematical Logic* (New York, 1940). Revised edition (Cambridge, Mass., 1951).
——, 'Notes on the Theory of Reference', in Quine, *From a Logical Point of View*, 130–38.
——, *From a Logical Point of View* (New York, 1953).
——, 'Speaking of Objects', *Proceedings and Addresses of the American Philosophical Association*, xxxi (1958), 5–22.
——, *Word and Object* (Cambridge, Mass., 1960).
——, 'Variables Explained Away', *Proceedings of the American Philosophical Society*, civ (1960), 343–7. Reprinted in Quine, *Selected Logic Papers*.
——, *Set Theory and Its Logic* (Cambridge, Mass., 1963). Revised edition (Cambridge, Mass., 1969).

——, *Selected Logic Papers* (New York, 1966).
——, 'Ontological Relativity', *The Journal of Philosophy*, lxv (1968), 185–212.
——, 'On the Reasons for the Indeterminacy of Translation', *The Journal of Philosophy*, lxvii (1970), 178–83.
Russell, B., 'On Denoting', *Mind*, xiv (1905), 479–93. Reprinted in Russell, *Logic and Knowledge*.
——, 'Mathematical Logic as Based on the Theory of Types', *American Journal of Mathematics*, xxx (1908), 222–62. Reprinted in Russell, *Logic and Knowledge*.
——, 'Knowledge By Acquaintance and Knowledge By Description', *Proceedings of the Aristotelian Society*, new series, xi (1910–11), 108–28. Reprinted in Russell, *Mysticism and Logic*.
——, *The Problems of Philosophy* (London, 1912). Paperback (London, 1959).
——, *Mysticism and Logic, and Other Essays* (London, 1918).
——, 'The Philosophy of Logical Atomism', *The Monist*, xxviii (1918), 495–527. Reprinted with revisions in Russell, *Logic and Knowledge*.
——, *Logic and Knowledge*, ed. R. C. Marsh (London, 1956).
——, and Whitehead, A. N., see Whitehead.
Savage, C. W., ed., *Perception and Cognition: Issues in the Foundations of Psychology* (Minnesota Studies in the Philosophy of Science, vol. ix), Minneapolis, 1978.
Sellars, W., and Chisholm, R., 'Intentionality and the Mental', in Feigl, Scriven, and Maxwell, 507–39.
Scheffler, I., 'An Inscriptional Approach to Indirect Discourse', *Analysis*, xiv (1954), 83–90.
Schiffer, S., *Meaning* (Oxford, 1972).
Schilpp, P., ed., *The Philosophy of Bertrand Russell* (The Library of Living Philosophers, vol. iii), New York, 1944.
Scott, D., 'More on the Axiom of Extensionality', in Bar-Hillel, Poznanski, Rabin, and Robinson, 115–31.
Shoemaker, S., 'Causality and Properties', to appear in van Inwagen.
Steiner, M., *Mathematical Knowledge* (Ithaca, 1975).
Strawson, P. F., 'On Referring', *Mind*, lix (1950), 320–44.
——, *Individuals* (London, 1959).
Tarski, A., 'Les fondements de la géométrie des corps', *Księga Pamiątkowa Pierwszego Polskiego Zjazdu Matematycznego*, supplement to *Annales de la Societé Polonaise de Mathématique* (Kraków, 1929), 29–33. Translated with additions in Tarski, *Logic, Semantics, Metamathematics*.
——, 'The Semantic Conception of Truth and the Foundations of Semantics', *Philosophy and Phenomenological Research*, iv (1944), 341–76. Reprinted in Linsky.
——, *Logic, Semantics, Metamathematics*, trans. J. H. Woodger (Oxford, 1956).
van Inwagen, P., see Inwagen.

Vermazen, B., 'Semantics and Semantics', *Foundations of Language*, vii (1971), 539–55.

Whitehead, A. N., and Russell, B., *Principia Mathematica*, vol. 1 (Cambridge, 1910).

Wittgenstein, L., *Tractatus Logico-Philosophicus* (London, 1922). English translation by C. K. Ogden of the German text of *Logisch-philosophische Abhandlung*. Republished with a new translation by D. F. Pears and B. F. McGuiness (London, 1961).

Yourgrau, W. and Breck, A. D., eds., *Physics, Logic and History: Proceedings of the 1966 Denver International Colloquium* (New York, 1970).

Symbols

Index

Numerals in parentheses refer to note numbers.

The Clarendon Library of Logic and Philosophy brings together books, by new as well as established authors, that combine originality of theme with rigour of statement. Its aim is to encourage new research of a professional standard into problems that are of current or perennial interest.

SOME BOOKS PUBLISHED IN THE SERIES

Ontological Economy: Substitutional Quantification and Mathematics
Dale Gottlieb

The Scientific Image
Bas C. van Fraassen
Also available in paperback

The Cement of the Universe: A Study of Causation
J. L. Mackie
Also available in paperback
'*The Cement of the Universe* is a major work on a central philosophical theme. It is impressive in scope as well as depth. All key issues that pertain to the nature of causation are dealt with.'
Bernard Berofsky in *The Journal of Philosophy*

Knowledge
Keith Lehrer
'This is a provocative and tightly argued essay in epistemology which is bound to become a classic . . . One of the best books in the theory of knowledge.'
David B. Annis in *Philosophia*

ISBN 0-19-824428-2
90000
9 780198 244288